THEORY OF COMPUTING, Volume 4 (2008), pp. 21–51
http://theoryofcomputing.org

Optimal lower bounds for the Korkine-Zolotareff parameters of a lattice and for Schnorr's algorithm for the shortest vector problem *

Miklós Ajtai

Received: November 15, 2006; published: May 10, 2008.

Abstract: Schnorr's algorithm for finding an approximation for the shortest nonzero vector in an n-dimensional lattice depends on a parameter k. He proved that for a fixed $k \leq n$ his algorithm (block $2k$-reduction) provides a lattice vector whose length is greater than the length of a shortest nonzero vector in the lattice by at most a factor of $(2k)^{2n/k}$. (The time required by the algorithm depends on k.) We show that if $k = o(n)$, this bound on the performance of Schnorr's algorithm cannot be improved (apart from a constant factor in the exponent). Namely, we prove the existence of a basis in $\mathbb{R}^n$ which is KZ-reduced on all k-segments and where the ratio $\|b_1\|/\mathtt{shortest}(L)$ is at least $k^{cn/k}$. Noting that such a basis renders all versions of Schnorr's algorithm idle (output = input), it follows that the quantity $k^{cn/k}$ is a lower bound on the approximation ratio any version of Schnorr's algorithm can achieve on the shortest vector problem. This proves that Schnorr's analysis of

*A preliminary version of this paper has appeared in the Proc. 35th ACM Symp. on Theory of Computing [2].

ACM Classification: F.2.2, G.2

AMS Classification: 68Q25, 68W40, 68W25, 11H55, 11H99, 52C07, 60D05

Key words and phrases: Approximation algorithm, lattice, shortest vector problem, geometry of numbers, basis reduction, LLL reduction, Schnorr's algorithm, Korkine-Zolotarev basis, random lattice

DOI: 10.4086/toc.2008.v004a002

the approximation ratio of his algorithm is optimal apart from the constant in the exponent. We also solve an open problem formulated by Schnorr about the Korkine-Zolotareff lattice constants α_k. We show that his upper bound $\alpha_k \le k^{1+\ln k}$ is the best possible apart from a constant factor in the exponent. We prove a similar result about his upper bound $\beta_k \le 4k^2$, where β_k is another lattice constant with an important role in Schnorr's analysis of his algorithm.

1 Introduction

1.1 Historical background, related results

One of the most important tasks of the algorithmic theory of lattices is to find a short nonzero vector in a given lattice. Although there is no known polynomial time algorithm which finds a shortest nonzero vector in an n-dimensional lattice, Lovász's algorithm also known as LLL reduction published in a paper of A. Lenstra, H. Lenstra, L. Lovász [12], finds a vector which is longer than the shortest vector by a factor of at most $2^{n-1/2}$. In this algorithm we repeatedly have to find short bases in two-dimensional lattices. The two-dimensional problem was already solved by Gauss [6], (see also Lagrange [11]). In 1987 C. P. Schnorr gave an algorithm which is a generalization of the LLL reduction (see [14]). In this algorithm we have to find short bases in $2k$-dimensional lattices. If we want a polynomial time algorithm then we may use the maximal k where a short basis can be found in polynomial time. Schnorr has proved that if k is fixed then the approximation factor in his algorithm is at most $(2k)^{2n/k}$.

We will prove that Schnorr's upper bound about his algorithm cannot be improved apart from a constant factor in the exponent. The proof is based on the construction of a lattice L and a basis $b_1,\dots,b_n$ in it, with the property that Schnorr's algorithm, given $b_1,\dots,b_n$ as an input, immediately terminates and gives b_1 as an approximation of the shortest nonzero vector of L. For a more detailed formulation of this result we need the following two (well-known) definitions.

Definition 1.1.

1. Let $b_1,\dots,b_n$ be a basis of the lattice L. Applying the Gram-Schmidt orthogonalization to this basis we get a sequence of vectors $b_1^*,\dots,b_n^*$ with the property that b_i^*, $i=1,2,\dots,n$ are pairwise orthogonal, $b_1=b_1^*$ and for all $i=1,\dots,n$ we have $b_i^*=b_i-\sum_{j=1}^{i-1}\mu_{i,j}b_j^*$ where $\mu_{i,j}=(b_i\cdot b_j^*)\|b_j^*\|^{-2}$. If P_i is the projection of the n-dimensional space onto the subspace orthogonal to $b_1,\dots,b_{i-1}$ then $b_i^*=P_ib_i$. We call the basis $b_1,\dots,b_n$ size-reduced if $|\mu_{i,j}|\le 1/2$ for all $1\le j<i\le n$. It is important for us that given an arbitrary basis $b_1,\dots,b_n$ it is easy to construct a size-reduced basis $d_1,\dots,d_n$ so that $b_i^*=d_i^*$ for $i=1,\dots,n$.

2. Let $b_1,\dots,b_n$ be a basis of the n-dimensional lattice L. For each $i=1,\dots,n$ let P_i be the orthogonal projection of the n-dimensional space onto the subspace orthogonal to the vectors $b_1,\dots,b_{i-1}$. Let $b_i^*=P_ib_i$. $b_1,\dots,b_n$ is a Korkine-Zolotareff basis (or a Korkine-Zolotareff reduced basis, see [10]) if it is size-reduced, and b_i^* is a shortest nonzero vector in the lattice P_iL, for all $i=1,\dots,n$.

We show that (apart from a constant factor in the exponent) the mentioned upper bound, about the performance of Schnorr's algorithm, cannot be improved. Namely there is an $\varepsilon' > 0$ so that for every $k \leq \varepsilon' n$ there exists a lattice and size reduced basis, $b_1, \dots, b_n$ in it so that, for every $i = 1, \dots, n-k+1$, the vectors $P_i b_i, \dots, P_i b_{i+k-1}$ form a Korkine-Zolotareff basis in the subspace generated by them. Consequently, Schnorr's algorithm is inactive on $b_1, \dots, b_n$ and so the short vector produced by the algorithm is b_1. However the construction will show that the ratio of $\|b_1\|/\lambda_1(L)$, where $\lambda_1(L)$ is the length of the shortest nonzero vector in L, is at least $k^{cn/k}$. So starting from this basis the output of Schnorr's algorithm is not better than the proven upper bound (apart from the mentioned constant factor). R. Kannan [8] constructed a lattice with these properties for LLL reduction, that is, for the $k = 2$ case. His construction is explicit while our present construction for an arbitrary k is probabilistic. Still we keep several important properties of Kannan's example. Although we give a tight lower bound on the approximation provided by Schnorr's algorithm, still two questions remain open. (1) What is the largest k where a short basis, in the sense of Schnorr's algorithm, can be found in time polynomial in n? (2) It is possible that if the same lattice is presented with another basis then Schnorr's algorithm gives a better result. (The largest k for which there is a known polynomial time deterministic algorithms is $k = O(\log n/\log\log n)$ given by R. Kannan in [9], and for probabilistic algorithms it is $k = O(\log n)$ given by M. Ajtai, R. Kumar, and D. Sivakumar in [4]. Therefore the approximation factor in deterministic polynomial time version of Schnorr's algorithm is $\exp\left(O((\log\log n)^2(\log n)^{-1}n)\right)$ while in the probabilistic version it is $\exp\left(O((\log\log n)(\log n)^{-1}n)\right)$.)

In a recent paper [15], Schnorr describes a practical algorithm for finding short vectors in lattices which in practice outperforms the known algorithms (there is no proven upper bound guaranteeing the good performance). There are concepts about lattice bases which play a role in both [15] and the present paper. The two papers complement each other in the sense that one shows the positive the other the negative effect of a basis with random behavior on the performance of algorithms.

1.2 The main results

Our proof about the worst-case performance of Schnorr's algorithm is based on the following random construction of a k-dimensional lattice $L = L_{k,c}$.

Definition 1.2. Assume the random variables $\mu_{i,j}$, $1 \leq j \leq i-2 \leq k-2$ are independent and uniformly distributed over the interval $(-1/2, 1/2)$. We define a lower triangular $k \times k$ matrix

$$A^{(k,c)} = (A_{i,j})_{i=1,\dots,k,j=1,\dots,k}\,.$$

The definition of the entries is the following.

$$\begin{aligned} A_{i,i} &= k^{-ci/k}, && \text{for } i = 1,\dots,k, \\ A_{i,i-1} &= \frac{1}{2}A_{i-1,i-1} = \frac{1}{2}k^{-c(i-1)/k}, && \text{for } i = 2,\dots,k, \text{ and} \\ A_{i,j} &= \mu_{i,j}A_{j,j}, && \text{if } 1 \leq j \leq i-2 \leq k-2\,. \end{aligned}$$

This implies that the ith row of the matrix $(A_{i,j})$ is of the following form:

$$\mu_{i,1}k^{-c/k} \quad \dots \quad \mu_{i,i-2}k^{-c(i-2)/k} \quad \tfrac{1}{2}k^{-c(i-1)/k} \quad k^{-ci/k} \quad 0 \quad \dots \quad 0$$

$L_{k,c}$ will denote the lattice in $\mathbb{R}^k$ generated by the rows of $A^{(k,c)}$.

The following theorem describes the properties of the lattice $L_{k,c}$ which will be crucial in our proof about the worst-case behavior of Schnorr's algorithm.

Theorem 1.3. *For all sufficiently small $c > 0$ there is an $\alpha > 0$ so that for all sufficiently large positive integers k with a probability of at least $1 - \mathrm{e}^{-\alpha k}$ we have the following. Let a_i be the ith row of the matrix $A^{(k,c)}$. Then $a_1, \ldots, a_k$ is a Korkine-Zolotareff reduced basis of the lattice $L_{k,c}$. In particular the first row of this matrix is a nonzero shortest vector of the lattice, with length $A_{1,1} = k^{-c/k}$. Moreover for all $i = 2, \ldots, k$, if P_i is the projection of $\mathbb{R}^k$ to the subspace orthogonal to $a_1, \ldots, a_{i-1}$ then in the $(k-i)$-dimensional lattice $P_i L_{k,c}$, the vector $P_i a_i$ is a nonzero shortest vector with length $A_{i,i} = k^{-ci/k}$.*

The theorem implies that if we apply the Gram-Schmidt orthogonalization to the basis $a_1, \ldots, a_k$ of $L_{k,c}$ then we get a basis a_i^* (of $\mathbb{R}^k$) with $a_i^* = k^{-ci/k} e_i$, where e_i is the ith unit vector.

We will use the lattice $L_{k,c}$ in the way described below. We will construct (at random) an n-dimensional lattice given with a basis $b_1, \ldots, b_n$, so that for every segment $b_s, \ldots, b_{s+k-1}$ of this basis, if we take the orthogonal projections of the vectors $b_s, \ldots, b_{s+k-1}$ to the subspace generated by $b_1, \ldots, b_{s+k-1}$, then the resulting sequence $b'_s, \ldots, b'_{s+k-1}$ is like the basis $a_1, \ldots, a_k$ from Theorem 1.3 multiplied by a constant factor. This and the described properties of the basis $a_1, \ldots, a_k$ will make it possible to show that Schnorr's algorithm is inactive on the basis $b_1, \ldots, b_n$. (The length of a shortest nonzero vector in the lattice generated by $b_1, \ldots, b_n$ will be estimated through Minkowski's convex body theorem.)

The basis $b_1, \ldots, b_n$ has a very concise definition similar to the definition of the lattice $L_{k,c}$ (although for the proofs we will reformulate it into a longer but more natural definition.) Our construction will depend on a constant $c > 0$ and we show that if c is sufficiently small then the constructed lattice and basis meet our requirements. Assume now that a constant $c > 0$ is fixed. The ith unit vector in $\mathbb{R}^n$ will be denoted by e_i.

Definition 1.4. Let $n, k, k \le n$ be positive integers and let $c > 0$ be a real number and let $f_i = k^{c(n-i+1)/k} e_i$ for $i = 1, \ldots, n$. Suppose further that $\mu_{i,j}$, $i = 1, \ldots, n$, $j = 1, \ldots, i-2$ are mutually independent random variables, uniformly distributed over the interval $(-1/2, 1/2)$. Let $b_1 = f_1$ and for all $i = 2, \ldots, n$ let $b_i = f_i + (1/2) f_{i-1} + \sum_{j=1}^{i-2} \mu_{i,j} f_j$. It is easy to see that the vectors $b_1, \ldots, b_n$ are linearly independent. Let $\Lambda(n,k,c)$ be the lattice generated by $b_1, \ldots, b_n$. The basis $(b_1, \ldots, b_n)$ will be denoted by $B(n,k,c)$. If the values of n, k, c are fixed then $\Lambda(n,k,c)$ and $B(n,k,c)$ are random variables.

Clearly this definition implies that the $n \times n$ matrix whose rows are $f_1, \ldots, f_n$ is lower triangular, and its ith row is

$$\mu_{i,1} k^{cn/k} \ldots \ \mu_{i,i-2} k^{c(n-(i-2)+1)/k)} \quad \frac{1}{2} k^{c(n-(i-1)+1)/k} \quad k^{c(n-i+1)/k}\, 0 \ldots 0$$

As a consequence $\Lambda_{n,k,c} = k^{c(n+1)/k} L_{n,c}$ and we get the basis $b_1, \ldots, b_n$ of $\Lambda_{n,k,c}$ by multiplying the rows of the matrix $A^{(n,c)}$ by $k^{c(n+1)/k}$.

Theorem 1.5. *For all sufficiently small $c > 0$, there exists $\varepsilon' > 0$ so that if n is a sufficiently large integer, and k is an integer with $k \leq \varepsilon' n$ then the following holds. Assume that L, $(b_1, \ldots, b_n)$ are the corresponding random values of the random variables $\Lambda(n,k,c)$, $B(n,k,c)$. Then with a positive probability the following conditions are satisfied.*

(1). $\|b_1\| = \exp\left(c\frac{n}{k}\log k\right)$, $\lambda_1(L) \leq 2n^{1/2}\exp\left(c\frac{n+1}{2k}\log k\right)$, and so

$$\|b_1\|/\lambda_1(L) \geq \frac{1}{2}n^{-1/2}k^{c(n-1)/(2k)} \geq k^{cn/(3k)},$$

(2). for all $i = 1, \ldots, n-2k+1$, the vectors $P_i b_i, \ldots, P_i b_{i+2k-1}$ form a Korkine-Zolotareff basis of the lattice generated by them.

Corollary 1.6. *The quantity $k^{cn/(3k)}$ is a lower bound on the approximation ratio any version of Schnorr's algorithm can achieve on the shortest vector problem, hence Schnorr's upper bound on the approximation ratio of his algorithm is optimal apart from the value of the constant in the exponent.*

Proof. Take a basis satisfying the properties guaranteed in Theorem 1.5. Then all versions of Schnorr's algorithm are idle on this basis (they do not change the basis). Therefore the algorithm's guess at the shortest vector is b_1, which is a factor of $k^{cn/(3k)}$ longer than the shortest vector in the lattice by Theorem 1.5. □

Remark 1.7.

1. The theorem does not provide a lower bound on the "positive probability" so it does not give a way to construct a lattice and a basis with the properties described in the theorem. We will see, however, from the proof that for all $c_1 > 0$ there is a $c_2 > 0$ so that if $k \geq c_2 \log n$ then L and $(b_1, \ldots, b_n)$ meet the requirements of the theorem with a probability $p \geq 1 - n^{-c_1}$. For smaller values of k, p can be exponentially small, so in this case the theorem does not give a construction. We will see that even in this case there is a probabilistic construction for L and $(b_1, \ldots, b_n)$ with the required properties.

2. The theorem holds, e. g., with any $\varepsilon' < c/10$ (this is not the best upper bound on ε'), and the final inequality $n^{-1/2}k^{c(n-1)/(2k)} \geq (1/2)k^{cn/(4k)}$ follows from this inequality. However if ε' is much larger than c, then the final inequality does not hold since the factor $n^{-1/2}$ will dominate and so we will have $(1/2)n^{-1/2}k^{c(n-1)/(2k)} < 1$.

In this paper we also solve an open problem asked by Schnorr about the geometry of lattices, which is related to the analysis of his algorithm.

Definition 1.8. For all positive integers k, the Korkine-Zolotareff constant is $\alpha_k = \sup \|b_1\|^2/\|b_k^*\|^2$, where the supremum is taken over all k-dimensional lattices L and over all Korkine-Zolotareff bases $b_1, \ldots, b_k$ in L.

Schnorr proves that $\alpha_k \leq k^{1+\log k}$ and asks whether $\alpha_k = k^{O(1)}$. We show that actually his upper bound is optimal up to a constant factor in the exponent. Namely we prove the following.

Theorem 1.9. *There is an $\varepsilon > 0$ so that the value of the Korkine-Zolotareff constant α_k is at least $k^{\varepsilon \log k}$ for all $k = 1,2,\dots$*

We will prove this result in Section 4. It will be an immediate consequence of Theorem 4.2. We note that the proof of Theorem 1.5 and Theorem 1.9 are both based on random lattice constructions. The two random constructions are similar but not identical. In fact, the construction used in the proof of Theorem 1.9 requires sharper estimates and so that proof involves more work.

Another lattice constant β_k is even more important for the determination of the approximation factor provided by Schnorr's algorithm. β_k defined as

$$\beta_k = \sup\left\{ \left(\left(\prod_{i=1}^{k} \|b_i^*\|^2\right) \left(\prod_{i=k+1}^{2k} \|b_i^*\|\right)^{-2} \right)^{1/k} \right\}$$

where the supremum is taken over all $2k$-dimensional lattices L and over all Korkine-Zolotareff bases $b_1,\dots,b_{2k}$ of L. Schnorr gives the upper bound $\beta_k \le 4k^2$ and uses it in the analysis of his algorithm. We show that this upper bound is tight apart from a constant in the exponent, that is, there is an $\varepsilon > 0$ so that $\beta_k \ge k^{\varepsilon}$ for all $k = 1,2,\dots$.

To prove the lower bound on β_k we will use the random lattice $L = \Lambda(2k,2k,c)$ with a sufficiently small $c > 0$ and show that with a high probability the basis $B(2k,2k,c)$ is a Korkine-Zolotareff basis in L. (Since $\Lambda(k,k,c) = k^{c(k+1)/k} L_{k,c}$, Theorem 1.3 implies this statement with $2k$ in the place of k.) For the lower bound on α_k we use a modified form of the random lattice $\Lambda(k,k,c)$, namely we will have $\|f_i\| = \exp\left(-c(\log(k-i+1))^2\right)$ for $i = 1,\dots,k-c_1$, and $\|f_i\| = \exp\left(-c(\log c_1)^2\right)$ for $i = k-c_1+1,\dots,k$, where c_1 is an integer sufficiently large with respect to c (but it does not depend on k). Otherwise the definition remains unchanged and we prove that with high probability $B(k,k,c)$ is a Korkine-Zolotareff basis for the modified B.

1.3 Motivation

The main motivation for this work is that Schnorr's algorithm gives the best proven approximation of the shortest nonzero vector in an n-dimensional lattice. So it is important to know how good is its performance. Moreover, no algorithm is known outside the framework of LLL and Schnorr's algorithm which gives comparable results. Therefore a lower bound on the algorithm shows the hardness of the approximate shortest vector problem at least according to our present knowledge. This limit on our knowledge seems to be serious since the approximation factor of Schnorr's algorithm has not been improved since its publication in 1987, apart from the increase of the largest k that can be used in poly-time (see [4]), but this did not affect the overall structure of the n-dimensional algorithm. Another possible use of the lattices constructed in this work is that we may try to find a better algorithm for finding an approximation of the shortest nonzero vector in a lattice by attacking this problem in the case of the counterexamples. Our probabilistically constructed lattices may be very good from this point of view because there seems to be no easy way to find a shorter vector in them than the one produced by Schnorr's algorithm. Motivated by this we formulate an open problem. In this open problem we present a specific random lattice ($\Lambda(n,k,c)$ for some choice of k and c) with the shortest known nonzero vector in it (which is b_1) and ask for a polynomial time algorithm which finds a nonzero vector shorter (or much shorter) than b_1. We may expect that a solution for this problem will contain some new idea about lattice algorithms, while it seems easier to attack the approximate shortest vector problem in a specific lattice than in its generality. On the other hand if no solution for this problem will be found for a long time then the lattice (which may also be generated together with a known short vector) may be useful for cryptographic purposes.

Open problem. Give a probabilistic algorithm $\mathcal{A}$ and a positive integer c_1 so that for all positive integers t,s and for all sufficiently large integers n if $\mathcal{A}$ gets $c = 1/t$, $k = \lfloor s\log n \rfloor$, n, and a random value of $B(n,k,c) = (b_1,\ldots,b_n)$ as an input then with a probability of at least $1/2$ (for the randomizations of both B and $\mathcal{A}$) and in time n^{c_1} the algorithm $\mathcal{A}$ finds a vector $x \in \Lambda(n,k,c)$ so that $x \neq 0$ and

(i.) $\|x\| < \|b_1\|$, or (a stronger requirement)

(ii.) $\|x\| < \frac{1}{2}\|b_1\|$.

Remark 1.10. The choice of k in the open problem is based on the fact that currently the largest k so that a Korkine-Zolotareff basis can be found in probabilistic polynomial time is $k = O(\log n)$ (see [4]), and so Schnorr's algorithm can be used with block length k. Therefore our open problem can be solved by improving this bound and still using Schnorr's algorithm.

In the proof of Theorem 1.5 we represent our lattice elements by sequences of integers in a similar way as it was done by R. Kannan in [9]. In fact his estimate about the number of such sequences which correspond to a short lattice vector remain valid in our case and is used in our proof. A similar bound was also used by M. Furst and R. Kannan in [5].

1.4 The history of the problem with some technical details

The history of finding short vectors in lattices starts with the works of Lagrange and Gauss. Both of them considered the problem of finding a shortest nonzero vector in two-dimensional lattices. In the nineteenth century Hermite, Zolotareff and Minkowski considered the problem of n-dimensional lattices as part of the reduction theory of quadratic forms. Although a large number of theorems in the theory of lattices were dealing with the existence of short vectors in lattices (e.g. Minkowski's convex body theorem) these were mainly existence theorems without giving any efficient method for finding short vectors. In 1983 A. Lenstra, H. Lenstra and L. Lovász found the first polynomial time algorithm for factoring polynomials with rational coefficients. Their solution was based on an algorithm which finds a vector not longer than $2^{(n-1)/2}$ times the shortest nonzero vector. Since then this method, the LLL reduction, has been used for the solution of a large number of both theoretical and practical problems. For example it was used to disprove the Mertens conjecture [13] and to break several proposed cryptosystems (see e.g. [7]).

The LLL reduction starts with an arbitrary basis $b_1,\ldots,b_n$ of the lattice L and then it gradually "improves" the basis. In each step we replace two consecutive elements of the basis by two new elements. Roughly speaking the goal of these exchanges is to get a size reduced basis $b_1,\ldots,b_n$ where for each i in the two-dimensional lattice K generated by P_ib_i and P_ib_{i+1}, the shortest vector is P_ib_i and (P_ib_i, P_ib_{i+1})

is a size reduced basis of K. We try to reach this goal by picking an i where this is not true and replacing b_i, b_{i+1} by two other vectors so that they satisfy this condition. (In order to do this we have to find the shortest nonzero vector in K, extend it into a size-reduced basis of K; this way we have the new $P_i b_i, P_{i+1} b_i$, then find the corresponding new b_i and b_{i+1}. All of these steps can be done easily.) In picking i we give preference to those pairs where we are far away from our goal according to some reasonable measure. In a polynomial number of steps, although we will not necessarily reach a situation where each pair satisfies the condition, still we can ensure that each pair will be close to it. This will guarantee that the ratios $\|b_i^*\|/\|b_{i+1}^*\|$ remain below a constant bound, namely $\|b_i^*\|/\|b_{i+1}^*\| \le \sqrt{2}$ for all $i = 1, \dots, n$. It is easy to see that $\lambda_1(L)$, the length of the shortest nonzero vector, is bounded from below by $\min_i \|b_i^*\|$ and so $\|b_1\| \le 2^{(n-1)/2} \lambda_1(L)$. (Indeed if u is a shortest nonzero vector then we consider the sequence $P_i u$, $i = 1, \dots, n$. We have $\|u\| = \|P_1 u\| \ge \|P_2 u\| \ge \dots \ge \|P_n u\| = 0$. Let j be the smallest integer so that $\|P_j u\| = 0$. By the definition of b_{j^*} we have that $P_{j-1} u = k b_j^*$, where $k \ne 0$ is an integer and so $\|u\| \ge \|b_j^*\|$.)

C. P. Schnorr has improved the approximation factor of the algorithm by working with k consecutive basis elements $b_i, \dots, b_{i+k-1}$ instead of just 2. This generalization creates significant new problems. What should be our goal for the k-blocks of basis elements and what will play the role of the inequality $\|b_i^*\|/\|b_{i+1}^*\| \le \sqrt{2}$? Perhaps most important is the difficulty that we have to prove everything for k-dimensional lattices instead of two-dimensional lattices where the arising problems do not cause significant difficulties.

Schnorr gave the following answers to these questions. (We just give a rough simplified sketch of his algorithm.) The goal is that every block $b_i, \dots, b_{i+k-1}$ should have the property that $P_i b_i, \dots, P_i b_{i+k-1}$ Korkine-Zolotareff basis in the lattice generated by them. Therefore in each step we select somehow an i so that this does not hold and replace the k basis vectors $b_i, \dots, b_{i+k-1}$ by k new vectors so that we have now a Korkine-Zolotareff basis. (For this step we have to find a Korkine-Zolotareff basis in the k-dimensional lattice generated by $P_i b_i, \dots, P_i b_{i+k-1}$. For polynomial time algorithms as we mentioned already this limits the size of k to $O(\log n)$ or $O(\log n / \log\log n)$.) Again we cannot hope in polynomial time to reach a stage where every block of length k provides a Korkine-Zolotareff basis, but we may reach a stage where every block is close to it according to some reasonable measure. We may get an upper bound on $\|b_1\|/\lambda_1(L)$ essentially by getting an upper bound on $\|b_i^*\|/\|b_{i+k-1}^*\|$ for all $i = 1, \dots, n-k+1$. This leads to the task, of finding an upper bound on $\|d_1\|^2/\|d_k^*\|^2$ for every k-dimensional lattice J, where $d_1, \dots, d_k$ is a Korkine-Zolotareff basis in J. In other words we need the value of the Korkine-Zolotareff constant defined earlier. Schnorr gave the upper bound $\alpha_k \le k^{1+\log k}$ and from this he got an upper bound on $\|b_1\|/\lambda_1(L)$. However, with another method (that we describe shortly), he gave a better upper bound on $\|b_1\|/\lambda_1(L)$ which made likely that the truth is $\alpha_k \le k^{O(1)}$ (this would have provided the same upper bound on $\|b_1\|/\lambda_1(L)$). In this paper we show that surprisingly Schnorr's upper bound on α_k is tight, that is, there is an $\varepsilon > 0$ with $\alpha_k \ge k^{\varepsilon \log k}$ for all k.

The other method used by Schnorr which gives the better upper bound on $\|b_1\|/\lambda_1(L)$ is the following. Instead of considering blocks of length k let us consider blocks of basis elements of length $2k$, $b_i, \dots, b_{i+2k-1}$. (We may simplify the picture and speed up the algorithm if we consider only the case when $n = mk$ and $i = tk + 1$.) The algorithm will be the same but now instead of $\|b_i^*\|^2/\|b_{i+k-1}^*\|^2$ we are interested in the geometric mean of such ratios on an interval of length k. That is, we consider the quantity

$$\Big(\big(\prod_{j=i}^{i+k-1} \|b_j^*\|^2 \big) \big(\prod_{j=i+k}^{i+2k-1} \|b_j^*\| \big)^{-2} \Big)^{1/k} .$$

This quantity also can be used to give an upper bound on $\|b_1\|/\lambda_1(L)$. Again we need an upper bound on the ratio, that is, we need an upper bound for the corresponding ratio for every $2k$-dimensional lattice with every possible choice of a Korkine-Zolotareff basis. By our definition given earlier the smallest such upper bound is the lattice constant β_k. Schnorr proved that $\beta_k \leq 4k^2$ and this lead him to the upper bound $(2k)^{2n/k}$ on the approximation factor. In this paper we show that this upper bound is tight apart from a constant factor in the exponent, that is, there is an $\varepsilon > 0$ so that $\beta_k \geq k^{\varepsilon}$ for all k.

The lower bound on β_k only shows that the analysis given by Schnorr about his algorithm cannot be improved by improving the upper bound on β_k but does not prove in itself that the algorithm cannot perform always better. We prove this latter statement by constructing a lattice L and a basis $b_1, \dots, b_n$ so that applying Schnorr's algorithm to this basis the resulting approximation factor, that is, $\|b_1\|/\lambda_1(L)$ is only $k^{\varepsilon n/k}$ for some constant ε, that is, apart from the factor $\varepsilon > 0$ it is the same as Schnorr's upper bound. Our basis $b_1, \dots, b_n$ will have the property that it is size-reduced and for any consecutive block of $2k$ basis vectors $b_i, \dots, b_{i+2k-1}$, the vectors $P_i b_i, \dots, P_i b_{i+2k-1}$ form a Korkine-Zolotareff basis of the lattice generated by them and

$$\left(\left(\prod_{j=i}^{i+k-1} \|P_i b_j^*\|^2 \right) \left(\prod_{j=i+k}^{i+2k-1} \|P_i b_j^*\| \right)^{-2} \right)^{1/k} \geq k^{\varepsilon}$$

where $\varepsilon > 0$ is a constant.

Remark 1.11.

1. In Schnorr's paper [14] several different algorithms are presented. Our lower bounds are valid even for the versions, k-reduction and block $2k$-reduction, where there is no polynomial time limit on the running time of the algorithm. The reason is that the basis that we provide for the lower bound has the property that starting from it Schnorr's algorithm immediately terminates.

2. A random lattice construction has been used to create problems with worst-case/average-case equivalence [3]. Another random lattice construction lead to a conjectured $0-1$ law for lattice properties testable in polynomial time.

   The present way of randomizing lattices is different from both of these although the choice of randomization was motivated by the randomization method of [1].

2 Sketch of the proofs of Theorems 1.3 and 1.5

First we give an equivalent definition for the random lattice $\Lambda(n,k,c)$, which is longer but more natural and has a clear motivation.

We will construct a size reduced basis $(b_1,\dots,b_n)$ so that

$$b_i^* = f_i = \exp\left(c\frac{n-i+1}{k}\log k\right)e_i\,,$$

$i=1,\dots,n$. Assume that we have such a lattice. Minkowski's convex body theorem implies that $2n^{1/2}(\det L)^{1/n}$ is an upper bound on the length of the shortest nonzero vector in any lattice. Now

$$(\det L)^{1/n} = \Big(\prod_{i=1}^{n}\|b_i^*\|\Big)^{1/n} = \exp\left(c\frac{n+1}{2k}\log k\right).$$

Since $\|b_1\| = \exp(c(n/k)\log k)$, we get (using that $k \le \varepsilon' n$ and that we may choose ε' with, e. g., $\varepsilon' < c/10$) that $\|b_1\|/\lambda_1(L) \ge k^{\varepsilon n/k}$. (This is the only point where we use the assumption $k \le \varepsilon' n$. For larger values of k we would need a better upper bound on $\lambda_1(L)$). This holds also if $(b_1,\dots,b_n) = B(n,2k,c)$. For the proof of Theorem 1.5 we have to show that $(P_ib_i,\dots,P_ib_{i+2k-1})$ is a Korkine-Zolotareff basis in the lattice generated by them for $i=1,\dots,n-2k+1$.

The vectors $b_1^*,\dots,b_n^*$ are already fixed and we have to define $b_1,\dots,b_n$. $b_1^*,\dots,b_n^*$ uniquely determine the projections P_i. Indeed P_i was defined as the orthogonal projection of $\mathbb{R}^n$ onto the subspace S orthogonal to $b_1,\dots,b_{i-1}$. The vectors $b_1,\dots,b_{i-1}$ generate the same subspace as the vectors $b_1^*,\dots,b_{i-1}^*$. Therefore S, and so P_i, are uniquely determined by $b_1^*,\dots,b_i^*$. We do not define the vectors b_i directly but define first the vectors P_jb_i for $j=n,\dots,1$ recursively by recursion on j starting at n and going downwards. (At the end we get b_i since $P_1b_i = b_i$.) This means that we build up the vectors b_i from their projected images by going backwards through the sequence of projections P_i. For $j=n$ we do not have any choices, by the definitions of b_i^* and P_j we have $P_nb_i = 0$ if $i<n$ and $P_nb_n = b_n^*$.

The basic principle of this process is the following. Assume that P_jb_i has been already defined for a fixed j and for all i. These elements generate an $n-j+1$-dimensional lattice L_j in the $n-j+1$-dimensional subspace generated by $b_j^*,\dots,b_n^*$. We will define a $(n-(j-1)+1)$-dimensional lattice L_{j-1} in the subspace generated by $b_{j-1}^*,\dots,b_n^*$ so that $P_jL_{j-1} = L_j$. Therefore if L_{j-1} is given somehow then $P_{j-1}b_i$ must be an element x of L_{j-1} with the property that $P_jx = P_jb_i$. Since we want our basis to be size reduced there are at most two possible choices for x since $P_{j-1}b_i$ and b_{j-1}^* has been already fixed. If in the definition of a size reduced basis we replace the requirement $|\mu_{i,j}| \le 1/2$ by $-1/2 < \mu_{i,j} \le 1/2$ then the choice of b_i is uniquely determined by the lattice L_{j-1}. Therefore our only task is the selection of the lattice L_{j-1} if we know already the lattice L_j. The conditions for this selection are the following: lattice L_{j-1} must be chosen from a given space whose dimension is larger by 1 than the dimension of L_j, it must contain a given vector b_{j-1}^* orthogonal to L_j, and we have to choose it in a way that a given orthogonal projection will map L_{j-1} onto L_j. There are infinitely many different choices for L_{j-1} with these properties. We will see that there is a very natural way to do it at random. However we will not do it completely at random since we want, that with a high probability, the vector $P_{j-1}b_{j-1} = b_{j-1}^*$ is no longer than the vector $P_{j-1}b_j$ (this is a consequence of the requirement that $P_{j-1}b_{j-1}$ is the first element of a Korkine-Zolotareff basis.) To make this requirement easily satisfiable we try to make $P_{j-1}b_j$ which is an inverse image of b_j^* as large as possible, otherwise we choose everything at random. Below we formulate this lattice selection problem in an abstract setting and define the randomization there.

We formulate a lemma below describing the random extension of an arbitrary m-dimensional lattice L. We will apply this lemma with $m := n-j+1$ and $L := L_j$. Assume now, for the formulation of the following lemma, that L is an arbitrary lattice in $\mathbb{R}^m$ and $a \in L$ so that a is contained in a basis of L. Assume further that $\kappa > 0$. Let P be the orthogonal projection P of $\mathbb{R}^{m+1}$ to $\mathbb{R}^m$ defined by $P(y_1,\dots,y_m,y_{m+1}) = (y_1,\dots,y_m)$. We are interested in lattices K in $\mathbb{R}^{m+1}$ so that $(0,\dots,0,\kappa) \in K$, $PK = L$, the shortest vector $w \in K$ with $Pw = a$ is as long as possible, otherwise K is random in some sense.

The assumption $(0,\dots,0,\kappa) \in K$ implies that for every K we have $\|w\|^2 \le \|a\|^2 + (\kappa/2)^2$ if $w \in K$ and $Pw = a$. However this upper bound is reached for the vector $(a, \kappa/2)$, therefore we will deal only with lattices K so that $(a,\kappa/2) \in K$. The following lemma describes a natural randomization of a lattice K with these properties. The definition of the randomization formally depends on an arbitrarily chosen basis of L containing a but the lemma states that the randomization is in fact independent from the choice of this basis.

Lemma 2.1. *Assume that m is a positive integer, $L \subseteq \mathbb{R}^m$ is a lattice, $a \in L$, $a \ne 0$, $\kappa > 0$ is a real number and $x_1,\dots,x_{m-1} \in L$ so that $\{x_1,\dots,x_{m-1},a\}$ is a basis of L. Depending on $m,\kappa,L,a,x_1,\dots,x_{m-1}$ we define a random variable Y whose values will be lattices in $\mathbb{R}^{m+1}$. We pick $m-1$ real numbers $\xi_1,\dots,\xi_{m-1}$ independently and with uniform distributions from the interval $(0,1)$. A random value K of Y will be the lattice in $\mathbb{R}^{m+1}$ generated by the vectors $(x_i,\xi_i\kappa)$, $i=1,\dots,m-1$, the vector $(a,\kappa/2)$ and the vector $(0,\dots,0,\kappa)$. Then the distribution of Y does not depend on the choice of the vectors $x_1,\dots,x_{m-1}$ and it does not change if we replace the vector a by the vector $-a$. Moreover for each possible value K of Y and for each $x \in L$ there is a unique real number σ_x, $-1/2 \le \sigma_x < 1/2$ so that $(x,\sigma_x\kappa) \in K$ and if x is linearly independent of the vector a then the distribution of σ_x is uniform over the interval $[-1/2,1/2)$.*

We return now to our basis construction. We will use Lemma 2.1 with the following substitutions: $m := n-j+1$; $L := L_j$; $a := b_j^*$; $\mathbb{R}^m$ is the subspace generated by $b_j^*,\dots,b_n^*$; $\mathbb{R}^{m+1}$ is the subspace generated by $b_{j-1},b_j^*,\dots,b_n^*$; $(0,\dots 0,\kappa) := b_{j-1}^*$.

Let K be the random lattice defined in the lemma. The property $(a,\kappa/2) \in K$ implies that $(1/2)b_{j-1}^* + b_j^* \in L_{j-1}$ and clearly $P_j((1/2)b_{j-1}^* + b_j^*) = b_j^*$. The other elements of L_j depend on the random choices defined by the random variable Y of the lemma.

As we have told, if L_{j-1} is given, the choice for $P_{j-1}b_i$ is unique with the condition $\mu_{i,j-1} \in (-1/2,1/2]$. This completes the definition of $P_{j-1}b_i$ for all i and so the definition of the basis $b_1,\dots,b_n$. It is an immediate consequence of the definition that $b_1,\dots,b_n$ are linearly independent and b_i^* is really the vector that we have selected in advance for this role. The condition $\mu_{i,j-1} \in (-1/2,1/2]$ guarantees that it is a size reduced basis. Lemma 2.1 implies that this definition is equivalent to our original definitions of $\Lambda(n,k,c)$, $B(n,k,c)$. Indeed, since we could pick the basis elements $\{x_i\}$ in an arbitrary way, using the basis $P_jb_j,\dots,P_jb_n$ gives the original definition.

We prove that with a positive probability for the randomization of $\Lambda(n,k,c)$ we have that for each $i=1,\dots,n-k+1$ the elements $P_ib_i,\dots,P_ib_{i+k-1}$ form a Korkine-Zolotareff basis of the lattice generated by them, provided that $c > 0$ is a sufficiently small constant. For this proof we do not need the whole n-dimensional lattice. The lattice construction can be described by restricting our attention to the elements of this k-dimensional lattice. In fact if we multiply every element of this lattice by a suitable real number (depending on n,k, and i) we get a lattice whose distribution is the same as the distribution of $\Lambda(k,k,c)$. We will show that $b_1,\dots,b_k$ is a Korkine-Zolotareff basis in $\Lambda(k,k,c)$ with a probability of at least $1-\mathrm{e}^{-\alpha k}$ where $\alpha > 0$ is a constant. This implies that there is a constant $c' > 0$ so that if $k > c'\log n$ then the probability that $P_ib_i, \ldots, P_ib_{i+k-1}$ is a Korkine-Zolotareff basis for all $i = 1, \ldots, n-k+1$, is close to 1. (For smaller integers k, this probability is not close to 1 but still it is not 0. For the proof of this fact we may either use Lovász's Local Lemma, or a more direct argument which also provides a construction.) This completes the sketch of the proof of Theorem 1.5.

3 The proofs of Theorems 1.3 and 1.5

First we prove Theorem 1.5 together with the lower bound on β_k, and then the lower bound on α_k. The proof about the lower bound on α_k has the same structure than the lower bound proof for β_k, but in places it requires different and somewhat sharper and more difficult estimates. In principle it would be possible to give the proof for β_k only and then point out the differences. However in this case, it would be very difficult to check the correctness of the second proof. Therefore we include the complete proof for α_k as well, although this includes some repetitions.

Definition 3.1. If n is a positive integer then the set $\{1, \ldots, n\}$ will be denoted by $\mathcal{N}_n$. $\mathcal{N}_0$ will be the empty set.

We will use the following observation in the proof of Lemma 2.1.

Lemma 3.2. *Suppose that $\xi_1, \ldots, \xi_k$ are independent random variables, uniformly distributed over the interval $(0,1)$. Assume further that $A = \{a_{i,j}\}$ is a k by k matrix with integer entries and with determinant ± 1. For each $i = 1, \ldots, k$ let $\rho_i = \sum_{j=1}^{k} a_{i,j}\xi_j$ and let $\nu_i = \rho_i - \lfloor \rho_i \rfloor$, where $\lfloor x \rfloor$ is the largest integer which is not greater than the real number x. Then each ν_i is uniformly distributed over $[0,1)$ and the random variables $\nu_1, \ldots, \nu_k$ are independent.*

Proof. It is sufficient to show that the vector $(\nu_1, \ldots, \nu_k)$ is uniformly distributed over the n-dimensional unit cube Q_k. The vector $(\rho_1, \ldots, \rho_k)$ is in the k-dimensional parallelepiped P which has a vertex at 0 and whose edges starting from 0 are the columns of A. The linear transformation A takes Q_k into P. Since the determinant of A is ± 1 we have that the distribution of the vector $\rho = (\rho_1, \ldots, \rho_k)$ is uniform over P. We get the vector $\nu = (\nu_1, \ldots, \nu_k)$ by reducing it modulo Q_k. Since both Q_k and P are basic parallelepipeds of the same lattice (the lattice of points with integer coordinates) this reduction is a one-to-one map which preserves the k-dimensional volume. Therefore the fact that ρ is uniformly distributed over P implies that ν is uniformly distributed over Q_k. □

Proof of Lemma 2.1. Since $a, x_1, \ldots, x_{m-1}$ are linearly independent and $\kappa \neq 0$ we have that the vectors $(x_i, \xi_i\kappa)$, $i = 1, \ldots, m-1$, $(a, \kappa/2)$ and the vector $(0, \ldots, 0, \kappa)$ are also linearly independent and so they form a basis of K. We will use this fact in the proof.

Let $z_1, \ldots, z_{m-1} \in L$ so that $z_1, \ldots, z_{m-1}, a$ is a basis of L. Assume that we take a random value K of Y using the basis $x_1, \ldots, x_{m-1}, a$. It is sufficient to show that with probability 1, for each $i = 1, \ldots, m-1$ there is a unique real number $\nu_i \in (0,1)$ so that $(z_i, \nu_i\kappa) \in K$, moreover the random variables ν_i, $i = 1, \ldots, k$ defined this way are uniformly distributed over $(0,1)$ and they are mutually independent.

First we consider the special case when there are integers t_i, $i = 1, \ldots, m-1$ so that $z_i = x_i + t_ia$, $i = 1, \ldots, m-1$. By the definition of ξ_i we have $(x_i, \xi_i\kappa) \in K$. This can be written in the form of $(z_i - t_ia, \xi_i\kappa) \in K$. Using that $(a, \kappa/2) \in K$ we get that $(z_i, (\xi_i + t_i/2)\kappa)) \in K$ and so $(0, \ldots, 0, \kappa) \in K$ implies the existence of ν_i. If ν_i is not unique then $(0,\dots,0,\vartheta\kappa)\in K$ for some $\vartheta\in(0,1)$, in contradiction to the fact that $(0,\dots,0,\kappa)$ is a basis vector of K. Since ν_i is the fractional part of $\xi_i+t_i/2$ and ξ_i is uniformly distributed over $(0,1)$, we have that ν_i is also uniformly distributed over $(0,1)$.

We consider now the general case assuming only that $(z_1,\dots,z_{m-1},a)$ is a basis of L. We have $z_i=t_ia_i+\sum_{j=1}^{m-1}\alpha_{i,j}x_j$ for $i=1,\dots,m-1$, where all of the coefficients t_i and $\alpha_{i,j}$ are integers. Let $w_i=z_i-t_ia$, $i=1,\dots,m-1$. Clearly $(w_1,\dots,w_{m-1},a)$ is a basis of L. It is sufficient to prove that our assertion holds with $z_i:=w_i$, $i=1,\dots,m-1$, since if this is true then we may apply the already proven special case with $x_i:=w_i$. In other words, we have to prove now the assertion for the special case when the basis vectors $z_1,\dots,z_{m-1}$ can be written in the form $z_i=\sum_{j=1}^{m-1}\alpha_{i,j}x_j$, $i=1,\dots,m-1$, where $\alpha_{i,j}$ is an integer for $i=1,\dots,m-1$, $j=1,\dots,m-1$. Since both $(z_1,\dots,z_{m-1},a)$ and $(x_1,\dots,x_m,a)$ are bases of L, we have that the determinant $|\alpha_{ij}|$ is ±1.

By the definition of K we have $(x_j,\xi_j\kappa)\in K$. For each fixed i we take the linear combinations of these vectors with coefficients $\alpha_{i,j}$. We get that $(z_i,(\sum_{j=1}^{m-1}\alpha_{i,j}\xi_j)\kappa)\in K$. Let ν_i be the fractional part of $\sum_{j=1}^{m-1}\alpha_{i,j}\xi_j$. Clearly with a probability 1 we have $\nu_j\in(0,1)$. Moreover Lemma 3.2 implies that each ν_i is uniformly distributed over $(0,1)$ and the random variables $\nu_1,\dots,\nu_k$ are independent.

Let $x\in L$. We have $x=ta+\sum_{i=1}^{m-1}b_ix_i$, where $t,b_1,\dots,b_{m-1}$ are integers. Since $(x_i,\xi_i\kappa)\in K$, $i=1,\dots,m-1$ and $(a,\kappa/2)\in K$ we have that $(x,\sigma'\kappa)\in K$, where $\sigma'=t/2+\sum_{i=1}^{n-1}b_i\xi_i$. Let σ_x the unique element of the interval $[-1/2,1/2)$ so that for a suitable integer s we have $\sigma'+s=\sigma_x$. $(0,\dots,0,\kappa)\in K$ implies that $(x,\sigma_x\kappa)\in K$. Assume that contrary to our assertion σ_x is not unique. By taking the difference of the two vectors $(x,\sigma_x\kappa)$ for two different values of σ we get that there is a $\vartheta\in(0,1)$ with $(0,\dots,0,\vartheta\kappa)\in K$ which contradicts to the fact that $(0,\dots,0,\kappa)$ is a basis vector of K. □

Definition 3.3.

1. Assume that m is a positive integer, $L\subseteq\mathbb{R}^m$ is a lattice, $\kappa>0$ is a real number and $a\in L$, $a\neq0$. The random variable Y defined in Lemma 2.1 will be denoted by $\mathtt{ext}_{m,\kappa,L,a}$.

   We define a random variable $\mathtt{rand}_{n,h}$ whose values will be lattices in $\mathbb{R}^n$. The definition depends on two parameters: a positive integer n and a function h with the property domain$(h)\supseteq\{1,..,n\}$, (only the values of h on $\{1,\dots,n\}$ will be used in the definition, but notationally it will be more convenient to allow functions with larger domains). The values of h will be positive real numbers. Assume that n and h are given. We define the random variable $\mathtt{rand}_{n,h}$ by recursion on n. At the same time we will prove by induction on n that for each possible value L of $\mathtt{rand}_{n,h}$ there is a minimal positive real number $\nu(L)$ so that $(0,\dots,0,\nu(L))\in L$. $\mathtt{rand}_{1,h}$ will be always the one-dimensional lattice consisting of all of the real numbers $ih(1)$, where i is an integer. Clearly $\nu(L)$ is $h(1)$. Assume that the random variable $\mathtt{rand}_{n-1,h}$ has been defined with values in $\mathbb{R}^{n-1}$ for some $n>1$. Then first we take a random value L of $\mathtt{rand}_{n-1,h}$. Let $a=(0,\dots,0,\nu(L))\in L$. Now we randomize $\mathtt{ext}_{n-1,h(n),L,a}$. Its value is a lattice $K\subseteq\mathbb{R}^n$, this will be the value of the random variable $\mathtt{rand}_{n,h}$. By the definition of $\mathtt{ext}_{n,h(n),L,a}$ the vector $(0,\dots,0,h(n))$ is in K which proves the existence of $\nu(K)$.

2. If n is a positive integer then $\pi^{(i,n)}$ will denote the orthogonal projection of $\mathbb{R}^n$ onto $\mathbb{R}^i$ defined by $\pi^{(i,n)}(y_1,\dots,y_i,y_{i+1},\dots,y_n)=(y_1,\dots,y_i)$. $\pi^{(0,n)}$ will be that map of $\mathbb{R}^n$ into the zero-dimensional

space $\mathbb{R}^0$ (consisting only of the vector 0). If the value of n is clear from the context we will write $\pi^{(i)}$ for $\pi^{(i,n)}$.

3. We say that the lattice $L \subseteq \mathbb{R}^n$ is triangular if it has a basis $a_i = (\alpha_{i,1}, \ldots, \alpha_{i,n})$, $i = 1, \ldots, n$ so that for all $i = 1, \ldots, n$, $\alpha_{i,i} \neq 0$ and for all $j = 1, \ldots, i-1$ we have $\alpha_{i,j} = 0$. (The condition $\alpha_{i,i} \neq 0$ can be omitted since for a basis it follows from the other conditions). We will say that the basis $a_1, \ldots, a_n$ is a triangular basis of L.

   We can formulate the definition of triangularity in the following equivalent form: L is triangular iff for each $i = 1, \ldots, n$ there is a real number $\alpha \neq 0$ so that $(0, \ldots, 0, \alpha) \in \pi^{(i)}(L)$.

Remark 3.4. The triangularity of the lattice L depends on the way it is embedded in $\mathbb{R}^n$. In particular this property is *not* invariant under isometric isomorphisms of lattices. Namely there is a lattice $L \subseteq \mathbb{R}^n$ and there is a unitary transformation U of $\mathbb{R}^n$ so that L is triangular but UL is not, where $UL = \{Ux \mid x \in L\}$.

The following lemma is an immediate consequence of the definitions.

Lemma 3.5. *Assume that n, i are positive integers, $i \leq n$, $L \subseteq \mathbb{R}^n$ is a triangular lattice, then $\pi^{(i)}L$ is a triangular lattice in $\mathbb{R}^i$.*

Lemma 3.6. *Assume that n is a positive integer, h is a function defined on the set of all positive integers whose values are positive real numbers. Then each value L of the random variable $\mathtt{rand}_{n,h}$ is a triangular lattice.*

Proof. The recursive definition of the random variable $\mathtt{rand}_{n,h}$ implies that for all $i = 1, \ldots, n$ we have that $(0, \ldots, 0, h(i)) \in \pi^{(i)}(L)$, which proves the triangularity of L. □

Suppose that $L \subseteq \mathbb{R}^n$ is a triangular lattice. Our next goal is to associate with each $x \in L$ a sequence of integers $a_1, \ldots, a_n$, which will behave in a similar way as sequence of coefficients in an orthogonal basis. In particular we will want to get a lower bound on $\|x\|^2$ using the sequence $a_1, \ldots, a_n$. We will derive this sequence by applying an integer version of the Gram-Schmidt orthogonalization to the triangular basis and following the projections of vector x during this orthogonalization. We will be able to achieve integrality since we will use lattice vectors during the whole process. These lattice vectors will not be necessarily in the lattice L. Triangularity implies that for each $i = 1, \ldots, n$, $\pi^{(i)}(L) \subseteq \mathbb{R}^i$ is an i-dimensional lattice. We will be able to use these lattices instead of L at the corresponding stages of the process. As a first step we define a sequence of vectors $\wp_{i,L} \in \pi^{(i)}(L)$, $i = 1, \ldots, n$ so that $\wp_{i,L}$ and e_i are parallel (in $\mathbb{R}^{(i)}$). We may consider $\wp_{i,L}$ as the result of the orthogonalization of the triangular basis.

Definition 3.7. Suppose that $L \subseteq \mathbb{R}^n$ is a triangular lattice. For each $i = 1, \ldots, n$ there is a $v = (v_0, \ldots, v_i) \in \pi^{(i)}L$ so that $v_1 = \cdots = v_{i-1} = 0$ and $v_i > 0$. v is unique up to a constant factor, so there is a unique vector v with this property if we add the requirement that v_i is minimal. We will denote this vector by $\wp_{i,L}$. Clearly every triangular basis of $\pi^{(i)}(L)$ contains $\wp_{i,L}$ or $-\wp_{i,L}$.

Remark 3.8. We will define a_i in the following way. We consider the set

$$S_x = \{\pi^{(i)}(x) + t\wp_{i,L} \mid t = 0, \pm 1, \pm 2, \ldots\}.$$

From the elements of S_x we take an $y_0 = \pi^{(i)}(x) + t_0\wp_{i,L}$ whose ith component is minimal in absolute value (preferring $1/2$ over $-1/2$). a_i will be defined by the equation $\pi^{(i)}(x) = y_0 + a_i\wp_{i,L}$.

The set S_x can be also defined as $S_x = \{y \in \pi^{(i)}(L) \mid \pi^{(i-1)}(y) = \pi^{(i-1)}(x)\}$. Then we get y_0 as a function of $\pi^{(i-1)}(x)$. Since we will need this function later, we will use it in the final definition of a_i given in Lemma 3.10 below.

Definition 3.9. If $L \subseteq \mathbb{R}^n$ is a triangular lattice, then for each $i \in [0, n-1]$ we define a function $\mathfrak{U}_{i,L}$ whose domain will be $\pi^{(i)}(L)$ and whose range will be in $\pi^{(i+1)}(L)$. We claim that for each $x \in \pi^{(i)}(L)$, there is a unique real number ξ in the interval $(-1/2, 1/2]$ so that $(x,0) + \xi\wp_{i+1,L} \in \pi^{(i+1)}(L)$. Indeed, since $x = \pi^{(i)}y$ for some $y \in L$ we have that $x' = \pi^{(i+1)}y \in \pi^{(i+1)}L$ and we get the vector x' by extending x with a new component. Since $\wp_{i+1,L}$ is contained in a triangular basis of $\pi^{(i+1)}(L)$ we can write x' in the form of $(x,0) + \xi\wp_{i+1,L}$. The uniqueness of ξ follows also from the fact that $\wp_{i_1,L}$ is contained in a triangular basis. We define a function $\mathfrak{U}_{i,L}$ mapping $\pi^{(i)}L$ into $\pi^{(i+1)}$ by $\mathfrak{U}_{i,L}(x) = (x,0) + \xi\wp_{i+1,L}$ where ξ is the unique real number in $(-1/2, 1/2]$ so that $(x,0) + \xi\wp_{i+1,L} \in \pi^{(i+1)}(L)$.

Lemma 3.10. *Assume that n is a positive integer, $L \subseteq \mathbb{R}^n$ is a triangular lattice. Then for all $x \in L$ there is a uniquely determined sequence of integers $a_1, \ldots, a_n$ so that for all $j = 1, \ldots, n$ we have*

$$\pi^{(j)}x = \mathfrak{U}_{j-1,L}(\pi^{(j-1)}x) + a_j\wp_{j,L}\,.$$

For this sequence $a_1, \ldots, a_n$ we have

$$\|x\|^2 \geq \sum_{j=1}^{n} \Big(\min\{|a_j|, \big||a_j| - \frac{1}{2}\big|\}\Big)^2 \|\wp_{j,L}\|^2 \geq \frac{1}{4}\sum_{j=1}^{n} a_j^2 \|\wp_{j,L}\|^2\,.$$

Proof. The vectors $\pi_j x$ and $\mathfrak{U}_{j-1,L}(\pi^{(j-1)}x)$ differ only in their jth coordinates. Since both are in $\pi^{(j)}L$ their difference $d = \pi_j x - \mathfrak{U}_{j-1,L}(\pi^{(j-1)}x)$ is a vector in $\pi^{(j)}L$ whose only nonzero coordinate is the jth one. $\wp_{i,L}$ is a basis vector of $\pi^{(j)}(L)$ therefore $d = a_j\wp_{j,L}$ for some integer a_j. Since $\wp_{j,L} \neq 0$, a_j is unique. Let $x = (x_1, \ldots, x_n)$. By the definition of $\mathfrak{U}_{j-1,L}(\pi^{(j-1)}x)$, the jth coordinate of $\mathfrak{U}_{j-1,L}(\pi^{(j-1)}x)$ can be written in the form of $\xi\wp_{j,L}$ where $\xi \in (-1/2, 1/2]$. Therefore $x_j = (\xi + a_j)\wp_{j,L}$. If ξ_j and a_j has identical signs or $a_j = 0$ then $|\xi + a_j| \geq |a_j|$. Otherwise $|\xi + a_j| \geq \big||a_j| - 1/2\big|$. This and the fact that the numbers a_i are integers imply the final sequence of inequalities. ☐

Definition 3.11.

1. If n is a positive integer, $L \subseteq \mathbb{R}^n$ is a triangular lattice, $1 \leq i < l \leq n$ are integers and $x \in L$ then the unique sequence $a = (a_1, \ldots, a_n)$ whose existence is guaranteed by Lemma 3.10 will be denoted $\vartheta^{(x,L)} = (\vartheta_1^{(x,L)}, \ldots, \vartheta_n^{(x,L)})$.

2. If $a = (a_1, \ldots, a_n)$ is a sequence of integers then `start`(a) will be the largest positive integer i so that $a_1 = \cdots = a_i = 0$. If there is no such positive integer then `start`$(a) = 0$. 4. Assume that L is a triangular lattice. For each $x \in L$ and $i = 1, \ldots, n$ let $y(i)$ be the unique element of L so that $\pi^{(i)}x = \pi^{(i)}y(i)$ and $\mathfrak{U}_{j,L}(\pi^{(j)}y(i)) = \pi^{(j+1)}y(i)$ for each $j = i, \ldots, n-1$. The uniqueness of $y(i)$ follows from the equality $y(i) = \mathfrak{U}_{n-1,L}(\ldots\mathfrak{U}_{i,L}(x)\ldots)$. We will denote the vector $y(i)$ by $\varphi(x,i)$.

3. If L is a triangular lattice then we define a basis $\mathbf{b}_{i,L}$, $i = 1,\dots,n$ in the following way. $b_{1,L} = \wp_{n,L}$ and for all $i = 2,\dots,n-1$, $\mathbf{b}_{i,L} = \varphi(\wp_{n-i+1}, n-i+1)$.

4. In the following we will assume that $\mathbb{R}^0 \subseteq \mathbb{R}^1 \subseteq \dots \subseteq \mathbb{R}^n$. Namely if $i < j$ we will identify the vector $(x_1,\dots,x_i) \in \mathbb{R}^i$ with the vector $(x_1,\dots,x_i,0,\dots,0) \in \mathbb{R}^j$. (A cleaner way to do this would have been to define $\mathbb{R}^n$ as the set of all of the infinite sequences $x_1,\dots,x_n,x_{n+1},\dots$ of real numbers so that $0 = x_{n+1} = x_{n+2} = \cdots$.)

Lemma 3.12. *Assume that L is a triangular lattice. Then the basis $\mathbf{b}_{i,L}$, $i = 1,\dots,n$ is size reduced.*

Proof. By the definition of $\mathbf{b}_{i,L}$ we have $\pi^{(n-i+1)}\mathbf{b}_{i,l} = \wp_{n-i+1,L}$. Since relative to the basis $\mathbf{b}_{i,L}$ we have $P_i = \pi^{(n-i+1)}$, where P_i is from the definition of the Gram-Schmidt orthogonalization, we have that $\mathbf{b}^*_{i,L} = \wp_{n-i+1,L}$. We have $\mathbf{b}_{i,L} = \mathfrak{U}_{n-1,L}(\dots\mathfrak{U}_{n-i+1,L}(\wp_{n-i+1,L})\dots)$. When we apply $\mathfrak{U}_{n-j,L}$ we get the $n-j+1$th coordinate of the vector $\mathbf{b}_{i,L}$ which is $\mu_{i,j}$. Therefore the definition of $\mathfrak{U}$ implies that $\mu_{i,j} \in (-1/2, 1/2]$ and so $|\mu_{i,j}| \leq 1/2$. □

Definition 3.13.

1. Let $\alpha_k = \sup |b_1|^2/|b_k^*|^2$, where the supremum is taken over all k-dimensional lattices L and over all Korkine-Zolotareff bases $b_1,\dots,b_k$ in L. The quantity α_k is the Korkine-Zolotareff constant.

2. We define another lattice constant β_k which depends on lattices of dimension $2k$.

$$\beta_k = \sup \left\{ \left(\Big(\prod_{i=1}^{k} \|b_i^*\|^2\Big) \Big(\prod_{i=k+1}^{2k} \|b_i^*\|\Big)^{-2} \right)^{1/k} \right\}$$

where the supremum is taken over all $2k$-dimensional lattices L and over all Korkine-Zolotareff bases $b_1,\dots,b_{2k}$ of L.

In [14] Schnorr proves that $\alpha_k \leq k^{1+\log k}$ and asks whether $\alpha_k = k^{O(1)}$. In Theorem 4.2 we will show that actually Schnorr's upper bound is tight namely there is an $\varepsilon > 0$ so that $\alpha_k > k^{\varepsilon \log k}$ for all $k = 1,2,\dots$. In Theorem 3.14 we show that Schnorr's upper bound on β_k, $\beta_k \leq 4k^2$ is also the best possible apart from a constant factor in the exponent. Schnorr's estimate of the approximation factor of his algorithm depends on the an upper bound on β_k. Theorem 3.14 implies that this analysis cannot be improved by giving a better upper bound on β_k. This in itself does not imply a lower bound on the worst-case performance of the algorithm. However based on the probabilistic lattice construction in the proof of Theorem 3.14 we will be able to construct a lattice with a basis which shows that the approximation factor in the worst-case is the same as in the upper bound apart from a constant factor in the exponent (Theorem 3.22 and Theorem 3.23).

Theorem 3.14. *There is an $\alpha > 0$ so that if $c > 0$ is a sufficiently small real number, n is a positive integer, and L is a random value of the random variable $\mathtt{rand}_{n,h}$, where $h(x) = \exp((cx/n)\log n) = n^{cx/n}$, then the following holds: With a probability of at least $1 - \mathrm{e}^{-\alpha n}$ the sequence $b_i = \mathbf{b}_{i,L}$, $i = 1,\dots,n$ is a Korkine-Zolotareff reduced basis of the n-dimensional lattice L. Moreover the Gram-Schmidt orthogonalization* *procedure applied to the basis $b_1,\dots,b_n$ yields the orthogonal vectors $b_i^* = \wp_{n-i+1,L}$, $i=1,\dots,n$. If $n=2m$ where m is an integer then we have*

$$\Big(\big(\prod_{i=1}^{m}\|b_i^*\|^2\big)\big(\prod_{i=m+1}^{2m}\|b_i^*\|\big)^{-2}\Big)^{1/m} = n^{2c}\,.$$

This theorem clearly implies Theorem 1.3.

Proof of Theorem 3.14 In the proof we will use the following properties of the lattice L, where L is a random value of $\mathtt{rand}_{n,h}$.

Lemma 3.15. *Suppose that n is a positive integer, h is an arbitrary function on the set of positive integers with positive real values and L is a possible value of the random variable $\mathtt{rand}_{n,h}$. Then the following requirements are met.*

(3)**.** L is a triangular lattice and for all $i=1,\dots,n$, $\pi^{(i)}\wp_{i,L} = (0,\dots,0,h(i))$

(4)**.** $\mathcal{U}_i(\wp_{i,L}) = (0,\dots,0,h(i),\frac{1}{2}h(i+1))$ for all $i=1,\dots,n-1$

(5)**.** for all $i=1,\dots,n-1$ and for all integers t we have $\mathcal{U}_i(t\wp_{i,L}) = (0,\dots,0,th(i),\frac{1}{2}h(i+1))$

Remark 3.16. In this lemma the function h is not necessarily the same as the function defined in Theorem 3.14.

Proof. (3) The triangularity of L is an immediate consequence of the definitions of $\mathtt{rand}$ and $\mathtt{ext}$. $\wp_{i,L}$ takes the role of the vector $(0,\dots,0,\kappa)$ in the definition of $\mathtt{ext}_{i-1,\kappa,\bar{L},a}$ where in our case $\bar{L} = \pi^{(i-1)}(L)$ and $\kappa = h(i)$.

(4) By the definition of $\mathtt{rand}_{n,h}$ we have that $\pi^{(i+1)}(L) = \mathtt{ext}_{n,\kappa,L',a}$, where $\kappa = h(i+1)$, $L' = \pi^{(i)}(L)$, and $a = \wp_{i,L} = (0,\dots,0,h(i))$. This definition also implies that $(a,\kappa/2)\in L'$, that is,

$$\Big(0,\dots,0,h(i),\frac{1}{2}h(i+1)\Big) \in \pi^{(i+1)}(L)\,.$$

Therefore the definition of $\mathcal{U}_i$ and $\wp_{i+1,L} = (0,\dots,0,h(i+1))$ implies (4). (5) is an immediate consequence of (4). ☐

We return to the proof of Theorem 3.14. First we estimate the probability of the following event A: "$\wp_{n,L}$ is not a shortest nonzero vector in L." Let $a=(a_1,\dots,a_n)$ be a sequence of integers. We estimate the probability of the event A_a: "there is an $x\in L$, $x\neq 0$ so that $\|x\| < \|\wp_n\|$ and $\vartheta^{(x,L)} = a$." Let $\mathtt{start}(a) = i$. This implies that $i<n$ otherwise because of $a_1=\dots=a_n=0$ we would have $x=0$. The condition $\mathtt{start}(a)=0$ also implies $\pi^{(i)}x=0$, $\pi^{(i+1)}x\neq 0$. The latter inequality holds since otherwise we would have $\pi^{(i+1)}x = \mathcal{U}_i(0) + a_{i+1}\wp_{i+1,L} = 0$ and therefore $a_{i+1}\wp_{i+1,L}=0$ which is impossible since $a_{i+1}\neq 0$ and $\wp_{i+1,L}\neq 0$. We distinguish four cases depending on i.

In the proof we will use repeatedly the following simple fact

(6)**.** If $h(x) = \exp\Big(\frac{cx}{n}\log n\Big)$ then $h(n-n(\log n)^{-1}) = h(n)\mathrm{e}^{-c}$.

Case I. $i > n - n(\log n)^{-1}$ or $n < (5/4)^{1/(2c)}$. We show that in this case $P(A_a) = 0$ since there is no $x \in L$ with $\|x\| < \|\wp_n\|$ and $\vartheta^{x,L} = a$. Indeed, assume that there is an $x \in L$ with these properties. $\mathtt{start}(a) = i$ implies $a_1 = \cdots = a_i = 0$ and so according to the definition of $\vartheta^{(x,L)} = a$ we have $\pi^{(j)}(x) = 0$ for $j = 1, \ldots, i$. $a_{i+1} \neq 0$ therefore $\pi^{(i)}(x) = 0$ implies

$$\pi^{(i+1)}(x) = a_{i+1}\pi^{(i+1)}\wp_{i+1,L} = (0, \ldots, 0, a_{i+1}h(i+1)).$$

If $i = n-1$ this implies $\|x\| \geq \wp_{n,L}$ in contradiction to our assumption. Therefore $i < n-1$. (5) implies that

$$\pi^{(i+2)}(x) = \mathfrak{U}_{i+1}(a_{i+1}\pi^{(i+1)}\wp_{i,L}) = \left(0, \ldots, 0, a_{i+1}h(i+1), \frac{1}{2}h(i+2)\right).$$

Consequently

$$\|x\|^2 \geq \|\pi^{(i+2)}x\|^2 \geq (h(i+1))^2 + \frac{1}{2}(h(i+2))^2 = n^{2c\frac{i+1}{n}} + \frac{1}{2}n^{2c\frac{i+2}{n}} \geq \frac{5}{4}n^{2ci/n}.$$

So the condition $i > n - n(\log n)^{-1}$ implies that

$$\|x\|^2 \geq \frac{5}{4}n^{2c(n-n(\log n)^{-1})n^{-1}} = \frac{5}{4}\exp\left(2c(1-(\log n)^{-1})\log n\right) = \frac{5}{4}\mathrm{e}^{2c\log n}\mathrm{e}^{-2c} > \mathrm{e}^{2c\log n} = \|\wp_{n,L}\|^2$$

in contradiction to our assumption. We consider now the condition $n < (5/4)^{1/(2c)}$. We have

$$\log\|x\|^2 \geq \log\left(\frac{5}{4}n^{2ci/n}\right) = \log\frac{5}{4} + 2cin^{-1}\log n = 2c\log n\left(\frac{i}{n} + (2c\log n)^{-1}\log\frac{5}{4}\right).$$

To give a lower bound on this we use that $n < (\frac{5}{4})^{1/(2c)}$ implies $(2c\log n)^{-1}\log\frac{5}{4} > 1$. Therefore

$$\log\|x\|^2 > 2c\log n\left(\frac{i}{n} + 1\right) \geq 2c\log n \geq \log\|\wp_{n,L}\|^2.$$

Case II. $n - n(\log n)^{-1} \geq i$, $n > (5/4)^{1/(2c)}$ and $|S_a| > (n-i')/10$ where $i' = \max\{i, \lceil n/2 \rceil\}$ and $S_a = \{j \in [i', n] \mid a_j \neq 0\}$. We show that in this case $P(A_a) = 0$ since there is no $x \in L$ with $\vartheta^{x,L} = a$ and $\|x\| < \|\wp_n\|$. Indeed assume that there is an $x \in L$ with these properties. By Lemma 3.10 we have

$$\begin{aligned}\|x\|^2 \geq \frac{1}{4}\sum\{\|\wp_j\|^2 \mid a_j \neq 0, j \in [i', n]\} \geq \frac{1}{4}|S_x|\|\wp_{i'}\|^2 \geq \frac{1}{4}|S_x|(h(i'))^2 >\\ \frac{1}{4}\frac{1}{10}(n-i')\exp\left(2c\frac{i'}{n}\log n\right) = \frac{1}{40}(n-i')\exp(2c\log n)\exp\left(-2c\frac{n-i'}{n}\log n\right) \geq\\ \frac{1}{40}(n-i')\exp(2c\log n)\exp\left(-2c\frac{n/2}{n}\log n\right) \geq \frac{1}{40}(n-n(\log n)^{-1})\mathrm{e}^{-c\log n}\|\wp_{n,L}^2\|.\end{aligned}$$

Since $n > (5/4)^{1/(2c)}$ and c is sufficiently small we have $\frac{1}{40}(n - n(\log n)^{-1})\mathrm{e}^{-c\log n} \geq 1$ and therefore $\|x\| \geq \|\wp_{n,L}\|$ in contradiction to our indirect assumption.

Case III. $n - n(\log n)^{-1} \geq i$, $n > (5/4)^{1/(2c)}$, $|S_a| \leq (1/10)(n-i')$ and $i > \lfloor n/2 \rfloor$. (The last inequality implies $i' = i$.) For each fixed $i > \lfloor n/2 \rfloor$, we estimate the probability of the event $A_a \wedge \mathtt{start}(a) = i$.

Suppose that L and a are fixed so that there is an $x \in L$ with $\vartheta^{(x,L)} = a$, $i = \mathtt{start}(a)$. By the definition of $\vartheta^{(x,L)}$, x is unique with this property. Let $Q = [i+2, n-1] \backslash S_a$. If $j \in Q$ then $\vartheta_j^{(x,L)} = a_j = 0$ and so $\pi^{(j)}(x) = \mathcal{U}_{j-1,L}(\pi^{(j-1)}x)$ and therefore by the definition of $\mathcal{U}_{j-1,L}$ we have that $\pi^{(j)}(x) = \pi^{(j-1)}(x) + \rho_j \wp_j$ for a uniquely defined $\rho_j \in (-1/2, 1/2]$. Let $M = h(n)(h(i))^{-1}$ and let A'_a be the event that "the number of integers $j \in Q$ with $|\rho_j| < (n-i)^{-1/5}M$ is at least $(n-i)/2$." We show that A_a implies A'_a and therefore $P(A_a) \leq P(A'_a)$, then we will prove an upper bound on $P(A')$. Assume that $\neg A'_a$. Then there are at least $3(n-i)/10$ integers $j \in Q$ so that $|\rho_j| \geq (n-i)^{-1/5}M$. Therefore

$$\|x\|^2 \geq \sum_{j \in Q} \rho_j^2 \|\wp_j\|^2 \geq \frac{3}{10}(n-i)(n-i)^{-2/5}M^2(h(i))^2 \geq$$

$$\frac{3}{10}(n-i)^{3/5}(h(n)(h(i))^{-1})^2(h(i))^2 = \frac{3}{10}(n-i)^{3/5}(h(n))^2 > \|\wp_{n,L}\|^2 ,$$

that is, $\neg A_a$ holds, which completes the proof of the implication $A'_a \Rightarrow A_a$.

Now we estimate $P(A'_a)$. Suppose $j \in Q$. ρ_j was defined by the conditions

$$\pi^{(j)}(x) = \pi^{(j-1)}(x) + \rho_j \wp_j ,$$

$\rho_j \in (-1/2, 1/2]$. Therefore the definitions of $\mathtt{rand}$, $\mathtt{ext}$ and Lemma 2.1 implies that if $\pi^{(j-1)}(x)$ is linearly independent of $\wp_{j-1}$ then ρ_j is uniformly distributed over $[-1/2, 1/2]$. However if $\pi^{(j-1)}(x) \neq 0$ and $\wp_{j-1}$ are linearly dependent then by (3) $\mathtt{start}(a) = j-1$. Since $i = \mathtt{start}(a)$, $i \leq j$, $Q \subseteq [i+2, n]$ this contradicts to $j \in Q$. Therefore the distribution of ρ_j is uniform over $[-1/2, 1/2]$. The definition of $\mathtt{rand}$ implies that random variables ρ_j, $j \in Q$ are mutually independent. Therefore the probability that "the number of integers $j \in Q$ with $|\rho_j| < (n-i)^{-1/5}M$ is at least $(n-i)/2$" is no more than $2^{n-i}(2(n-i)^{-1/5}M)^{1/2(n-i)}$. We get that

(7). $P(A_a) \leq P(A'_a) \leq 2^{n-i}(2(n-i)^{-1/5}M)^{1/2(n-i)}$.

Definition 3.17. Assume that n is a positive integer h is an arbitrary function defined on the set of positive integers with positive real values, L is a possible value of the random variable $\mathtt{rand}_{n,h}$ and $a = (a_1, \ldots, a_n)$ is a sequence of integers. We say that the sequence a is acceptable if there is at least one possible value L of the random variable $\mathtt{rand}_{n,h}$ and an $x \in L$ with $a = \vartheta^{x,L}$ and $\|x\| < \|\wp_n\|$.

The following lemma is an estimate on the number of acceptable sequences. It was formulated and proved, in a somewhat different context and with different terminology, by R. Kannan in [9]. These differences however do not affect the validity of the upper bound and its proof.

Lemma 3.18 (Kannan). *Assume that $c > 0$ and n, i, $i \leq n$ are positive integers, h is an arbitrary function defined on the set of positive integers with positive real values, L is a possible value of the random variable $\mathtt{rand}_{n,h}$ and $a = (a_1, \ldots, a_n)$ is an acceptable sequence with $\mathtt{start}(a) = i$. Then for all $j = i, \ldots, n$ we have $|a_j| < h(n)(h(j))^{-1}$. Moreover the number of acceptable sequences with $\mathtt{start}(a) \geq i$ is at most*

$$\prod_{j=i}^{n}(1 + 2h(n)(h(j))^{-1}) \leq 3^{n-i}\prod_{j=i}^{n}(h(n)(h(j))^{-1}) .$$

Remark 3.19. In this lemma the function h is not necessarily the same as in Theorem 3.14.

Proof of Lemma 3.18. Assume that a is an acceptable sequence and let L be a value of $\mathtt{rand}_{n,h}$ where there is an $x \in L$ with $\|x\| \le \wp_{n,L}$ and $\vartheta^{(x,L)} = a$. According to the definition of a_j we have $\pi^{(j)}(x) = (\mathfrak{U}_{j-1,L}(\pi^{(j-1)}x)) + a_j \wp_{j,L}$. Therefore

$$h(n) = \|\wp_{n,L}\| > \|x\| \ge \|\pi^{(j)}x\| \ge |a_j|\|\wp_j\| = |a_j|h(j)$$

which implies the inequality $|a_j| < h(n)(h(j))^{-1}$.

If a is an acceptable sequence with $\mathtt{start}(a) \ge i$ then $a_j = 0$ for all $j = 1,\dots,i$. By the inequality proven above, for each $j > i$ the number of choices for a_j is at most $1 + 2h(n)(h(j))^{-1}$. ☐

We continue now the proof of Case III. For a fixed $i > n/2$ let p_i be the probability of the following event: there is an acceptable sequence a so that A_a. Clearly

$$p_i \le \sum \{P(A_a) \mid a \text{ is acceptable and } i = \mathtt{start}(a)\}.$$

Using our upper bound (7) on $P(A_a)$ and the upper bound in Lemma 3.18 for the number of acceptable sequences we get

$$p_i \le 2^{n-i}(2(n-i)^{-1/5}M)^{(n-i)/2}3^{n-i}\prod_{j=i}^{n}\left(h(n)(h(j))^{-1}\right).$$

Using that for all $j \in [i,n]$ we have $h(n)(h(j))^{-1} \le h(n)(h(i))^{-1} = M$ we get that

$$p_i \le 3^{n-i}2^{(n-i)/2}M^{3(n-i)/2}(n-i)^{-(n-i)/10}.$$

We give an upper bound on M.

$$M = h(n)/h(i) = \exp\left(c\frac{n-i}{n}\log n\right) = \exp\left(c\frac{n-i}{n}(\log\frac{n}{n-i} + \log(n-i))\right).$$

Since the function $(1/x)\log x$ takes its maximum at $x = \mathrm{e}$ in the interval $[1,\infty)$, we have that $(1/x)\log x \le 1/\mathrm{e} < 1$ if $x \ge 1$. This yields

$$M \le \mathrm{e}^c \exp\left(c\frac{n-i}{n}\log(n-i)\right) \le \mathrm{e}^c\mathrm{e}^{c\log(n-i)} \le \mathrm{e}^c(n-i)^c.$$

Substituting this into the upper bound on p_i and using that c is sufficiently small we get $p_i \le (n-i)^{-(n-i)/20}$. Therefore the probability that there is an i so that A_a holds for a suitably chosen a is at most

$$\sum_{i=n/2}^{n-n(\log n)^{-1}} (n-i)^{-(n-i)/20}.$$

It is easy to see that in this sum the largest term (with $i = n - n(\log n)^{-1}$) is greater than the sum of the others. So we get

$$p_i \le 2\left(\frac{n}{\log n}\right)^{-1/20\frac{n}{\log n}} \le 2\exp\left(-\frac{1}{20}n - n\log\log n(\log n)^{-1}\right) \le \mathrm{e}^{-n/21}.$$

Case IV. $n > (5/4)^{1/(2c)}$, $|S_a| \le (n-i')/10$ and $i \le \lfloor n/2 \rfloor$. (The last inequality implies $i' = \lfloor n/2 \rfloor$.)

For each acceptable sequence a with $\texttt{start}(a) \le n/2$ we will give an upper bound on the probability of the event A_a. Then we will estimate the number of acceptable sequences with this property.

Suppose that L and a are fixed so that there is an $x \in L$ with $\vartheta^{(x,L)} = a$. We have

$$\pi^{(j)}(x) = (\mathcal{U}_{j-1,L}(\pi^{(j-1)})x) + a_j \wp_j$$

for all $j = 0, \dots, n-1$. Let $Q = [n/2, n-1] \setminus S_a$. If $j \in Q$ then

$$\vartheta_j^{(x,L)} = a_j = 0 \quad \text{and so} \quad \pi^{(j)}(x) = \mathcal{U}_{j-1,L}(\pi^{(j-1)}x)$$

and therefore by the definition of $\mathcal{U}_{j-1,L}$ we have that $\pi^{(j)}(x) = \pi^{(j-1)}(x) + \rho_j \wp_j$ for a uniquely defined $\rho_j \in (-1/2, 1/2]$. Let A'_a be the event that "the number of integers $j \in Q$ with $|\rho_j| < n^{-\tau}$ is at least $n/4$," where $\tau < 1/4$ is chosen so that c is sufficiently small with respect to τ. (This definition is somewhat different from the corresponding definition in Case III.) We show that A_a implies A'_a and therefore $P(A_a) \le P(A'_a)$; then we will prove the claimed upper bound on $P(A')$.

Assume that $\neg A'_a$. Then there are at least $n/5$ integers $j \in Q$ so that $|\rho_j| \ge n^{-\tau}$. Therefore

$$\begin{aligned}\|x\|^2 \ge \sum_{j \in Q} \rho_j^2 \|\wp_j\|^2 \ge \frac{1}{5} n n^{-2\tau} \left(\frac{hn}{2}\right)^2 = \frac{1}{5} n^{1-2\tau} \left(\frac{hn}{2}\right)^2 \\ = \frac{1}{5} n^{1-2\tau} \exp\left(-2c \frac{n-(n/2)}{n} \log n\right) = \frac{1}{5} n^{1-2\tau} n^{-c}.\end{aligned}$$

Since $\tau < 1/4$ and $c > 0$ is sufficiently small with respect to τ this implies that $\|x\|^2 > n^{2c} = (h(n))^2 = \wp_n^2$, that is, $\neg A_a$ holds, which completes the proof of the implication $A'_a \Rightarrow A_a$.

In Case III we have seen that the random variables ρ_j are independent and uniformly distributed over $[-1/2, 1/2]$. Therefore the probability that "the number of integers $j \in Q$ with $|\rho_j| < n^{-\tau}$ is at least $n/4$" is no more than $2^{n/2}(2n^{-\tau})^{n/4}$. We get that

(8). $P(A_a) \le P(A'_a) \le 2^{n/2}(2n^{-\tau})^{n/4}$.

Now we estimate the number of acceptable sequences with $\texttt{start}(a) \le n/2$. This is no more than the total number of acceptable sequences without any restrictions. By Lemma 3.18 this number is at most $\prod_{j=1}^n 3^n h(n)(h(j))^{-1}$. Therefore if p is the probability of the event that there is an acceptable sequence a so that $\texttt{start}(a) \le n/2$ and A_a holds then using our upper bounds we get

$$p \le 2^{n/2}(2n^{-\tau})^{n/4} 3^n \prod_{j=1}^{n} (h(n)(h(j))^{-1}).$$

Using that $h(j) = \exp((cj/n)\log n)$ we can calculate the value of the product exactly:

$$\prod_{j=1}^{n} (h(n)(h(j))^{-1}) = (\mathrm{e}^{c \log n})^n \exp\left(-c \frac{n(n+1)}{2n} \log n\right) = \exp\left(c \frac{n-1}{2} \log n\right).$$

Substituting this inequality in the upper bound on p we get:

$$\log p \le \frac{n}{2}\log 2 - \tau(\log n + \log 2)\frac{n}{4} + n\log 3 + c\frac{n-1}{2}\log n\,.$$

Since $c > 0$ is sufficiently small with respect to τ the term $-(\tau n/4)\log n$ will dominate in absolute value therefore $p \le n^{-\tau n/8}$. This is our upper bound on the probability that there is an $x \in L\backslash\{0, \pm\wp_{n,L}\}$ with $\|x\| \le \|\wp_{n,L}\|$.

Summarizing the results of the four different cases we get the following:

Lemma 3.20. *If L is the random lattice defined in Theorem 3.14, then the probability that $\wp_{n,L}$ is not a shortest nonzero vector of L is at most $3\mathrm{e}^{-n/21}$.*

Suppose that L is a random value of $\mathtt{rand}_{n,h}$. We show now that with high probability for all $i = 1,\dots,n$, $\wp_{i,n}$ is a shortest nonzero vector in $\pi^{(i)}(L)$. Indeed if we multiply every element of $\pi^{(i)}L$ by the number $h(n)(h(i))^{-1}$ we get an i-dimensional triangular lattice K. We have $\wp_{j,K} = h(n)(h(i))^{-1}\wp_{j,L}$ for all $j = 1,\dots,i$. Therefore

$$\begin{aligned}\|\wp_{j,K}\| = h(n)(h(i))^{-1}\|\wp_{j,L}\| &= \exp\left(c\frac{n}{n}\log n\right)\exp\left(-c\frac{i}{n}\log n\right)\exp\left(c\frac{j}{n}\log n\right)\\ &= \exp\left(c\frac{n-(i-j)}{n}\log n\right) = h(n-(i-j))\,.\end{aligned}$$

Therefore $\|\wp_{i,K}\| = h(n)$, $\|\wp_{i-1,K}\| = h(n-1),\dots,\|\wp_{1,K}\| = h(n-i+1)$. This and the definition of $\mathtt{rand}$ implies that the lattice K can be embedded into another triangular lattice J so that $\wp_{i,K} = \wp_{n,J}, \wp_{i-1,K} = \wp_{n-1,J},\dots,\wp_{1,K} = \wp_{1+(n-i),J}$ and J is a random value of $\mathtt{rand}_{n,h}$. Lemma 3.20 implies that with a probability of at least $1-3\mathrm{e}^{-n/21}$ we have that $\wp_{n,J} = \wp_{i,K} = h(n)(h(i))^{-1}\wp_{i,L}$ is a shortest nonzero vector in J. Suppose that for some outcome of the randomization it is a shortest nonzero vector in J. Since $K \subseteq J$, it is also a shortest nonzero vector in K and finally since $\pi^i(L) = h(n)^{-1}(h(i))K$, the vector $\wp_{i,L} = h(n)^{-1}(h(i))\wp_{i,K}$ is a shortest nonzero vector in $\pi^i(L)$ with a probability of at least $1-3\mathrm{e}^{-n/21}$. The probability that this happens for all $i = 1,\dots,n$ simultaneously is at least $1-3n\mathrm{e}^{-n/21} \ge 1-\mathrm{e}^{-n/22}$ if n is sufficiently large. The definition of the basis $b_i = \mathbf{b}_{i,L}$ implies that $\pi^{(i)}b_{n-i+1} = \wp_{i,L}$, therefore if we apply the Gram-Schmidt orthogonalization procedure to the basis b_i then we get $b_i^* = \wp_i$ and so the probability that b_i is a Korkine-Zolotareff basis is at least $1-\mathrm{e}^{-n/22}$. Now

$$\|\wp_{j,L}\| = h(j) = \exp\left(c\frac{j}{n}\log n\right)$$

implies the last equation of the Theorem.

This concludes the proof of Theorem 3.14 (and Theorem 1.3).

Definition 3.21. Assume that $B = (b_1,\dots,b_n)$ is a basis of the lattice L and $1 \le i \le n$. Then $P_i^{(B)}$ will denote the orthogonal projection of L onto the subspace orthogonal to $b_1,\dots,b_{i-1}$.

Theorem 3.22. *There exists an $\alpha > 0$, so that if $c > 0$ is sufficiently small, k,n are positive integers $2 \le k \le n$, and L is a random value of $\mathtt{rand}_{n,g}$ where $g(j) = \exp((cj/k)\log k)$, $b_i = \mathbf{b}_{i,L}$ for $i = 1,\dots,n$, and $B = (b_1,\dots,b_n)$, then the following holds. Assume that $p_{n,k}$ is the probability of the following event: "For all $i = 1,\dots,n-k$ the orthogonal projections of $P_i^{(B)} b_i,\dots,P_i^{(B)} b_{i+k-1}$ is a Korkine-Zolotareff basis of the k-dimensional lattice generated by these vectors." Then $p_{n,k} \neq 0$ and $p_{n,k} \ge 1-(n-k)\mathrm{e}^{-\alpha k}$.*

Proof. The second inequality is an immediate consequence of Theorem 3.14. Indeed for a fixed i let K be the lattice generated by $P_i^{(B)} b_i,\dots,P_i^{(B)} b_{i+k-1}$. Clearly $P_i^{(B)} b_i = \pi^{(i)} b_i$. Therefore

$$\exp\left(c\frac{k-i}{k}\log k\right) K$$

is a value of the random variable $\mathtt{rand}_{k,h}$ where $h(j) = \exp((cj/k)\log k)$. Therefore, by Theorem 3.14, the probability that $P_i^{(B)} b_i,\dots,P_i^{(B)} b_{i+k-1}$ is not a Korkine-Zolotareff basis is at most $\mathrm{e}^{-\alpha k}$. The fact that there are at most $n-k$ choices for i implies the second inequality in the conclusion of the theorem.

If n is sufficiently large with respect to k then the proven second inequality does not imply $p_{n,k} \neq 0$. Still it can be shown by the Lovász Local Lemma that it holds. Actually the present situation is a very simple special case of Lovász's lemma so a direct proof and a corresponding probabilistic construction can be easily given based on the following observation.

Let $0 < \varepsilon < 1/10$. Suppose that $Q(x,y)$ is a binary relation on the finite set A so that if x,y are chosen at random, independently and with uniform distribution from A then $P(Q(x,y)) > 1-\varepsilon$. For each $\tau > 0$ let A_τ be the set of all $x \in A$ with the property that if we take a random $y \in A$ with uniform distribution then $P(Q(x,y)) \ge 1-\tau$. Then $|A|^{-1}|A_\tau| \ge 1-\varepsilon/\tau$. Indeed if we count the number of pairs (x,y) with $\neg Q(x,y)$ for all $x \in A \backslash A_\tau$ together then we get that this number is at least $(|A|-|A_\tau|)\tau|A|$. On the other hand from our assumption we know that this number is at most $\varepsilon|A|^2$ which implies that claimed inequality.

If we pick now $\tau = \sqrt{\varepsilon}$ then we have $|A|^{-1}|A_\tau| \ge 1-\sqrt{\varepsilon}$. Therefore for each $x \in A_\tau$ the number of elements $y \in A_\tau$ with $Q(x,y)$ is at least $|A|-2\sqrt{\varepsilon}|A|$. This implies that if $x_0 \in A_\tau$ then we may pick a sequence of elements $x_0,x_1,x_2,\dots,x_m$ recursively (for an arbitrary m) so that for all $i = 1,\dots,m$, $x_i \in A_\tau$ and $Q(x_{i-1},x_i)$. This recursive construction can be turned into an algorithm if the validity of the relation $Q(x,y)$ can be algorithmically decided. This implies that although we may not be able to decide with high probability whether an element is in A_τ but we can tell about all of the elements of $A_{\tau/2}$ with high probability that they are in A_τ. Therefore picking random elements and testing them whether they are in A_τ (with high probability) we can build the sequence $x_0,x_1,\dots,x_m$.

This recursive procedure makes it possible to prove the existence of a basis with high probability and with the required properties even in the case when k is small compared to n. To get a construction we need that for dimension at most k we are able to verify whether a given vector is a nonzero shortest vector of the lattice. However if we are able to perform Schnorr's algorithm with parameter k, then we have already this ability. (By polynomial time computation we can find the shortest vector in dimension $O(\log n)$. This seems to match the lower bound for the k where with high probability we get Korkine-Zolotareff bases in all of the blocks of length k. However it is not clear whether the constant factors of $\log n$ in the bounds really overlap.) □

Theorem 3.23. *There exist $\varepsilon > 0, \varepsilon' > 0$ so that for all positive integers k, n, $k \le \varepsilon' n$, there is an n-dimensional lattice L and a basis $b_1, \dots, b_n$ in L so that if $b_1^*, \dots, b_n^*$ are the orthogonal vectors that we get through Gram-Schmidt orthogonalization from the basis $b_1, \dots, b_n$ then the following holds. For each $s = 1, \dots, n-2k-1$ we have*

$$\left(\Big(\prod_{i=s+1}^{k} \|b_i^*\|^2 \Big) \Big(\prod_{i=s+k+1}^{2k} \|b_i^*\| \Big)^{-2} \right)^{1/k} \le k^{\varepsilon \log k},$$

$\|b_1\|/\lambda_1(L) \ge k^{\varepsilon \frac{n}{k}}$, where $\lambda_1(L)$ is the length of the shortest nonzero vector in L.

Proof. L will be a lattice whose existence is guaranteed by Theorem 3.22 with $2k$ in the place of k. We have

$$\|b_i^*\| = g(n-i+1) = \exp\left(c \frac{n-i+1}{2k} \log 2k \right).$$

This implies the first inequality. According to Minkowski's convex body theorem we have

$$\lambda_1(L) \le n^{1/2} (\det(L))^{1/n} = n^{1/2} \left(\prod_{i=1}^{n} \|b_i^*\| \right)^{1/n}.$$

Using that and $k \le \varepsilon' n$ we get $\|b_1\|/\lambda_1(L) \ge k^{\varepsilon n/k}$. □

4 The lower bound on the Korkine-Zolotareff constant α_k

Definition 4.1. For each positive integer k and real number c we define a function $\mathcal{F}_{k,c}$ whose domain is the set of all positive integers. For all $i = 1, \dots, k$ let $\mathcal{F}_{k,c}(i) = \mathrm{e}^{c(\log k)^2}$ and for all $i = k, k+1, \dots$ let $\mathcal{F}_{k,c}(i) = \mathrm{e}^{c(\log i)^2}$.

Theorem 4.2. *Assume that $c > 0$ is a sufficiently small real number and $c_1 > 0$ is a sufficiently large positive integer. For all $n = 1, 2, \dots$ if L is a random value of the random variable $\mathtt{rand}_{n,\mathcal{F}_{c_1,c}}$ then the following holds. With a probability of at least $1/2$ the sequence $b_i = \mathbf{b}_{i,L}$, $i = 1, \dots, n$ is a Korkine-Zolotareff reduced basis of the n-dimensional lattice L. Moreover the Gram-Schmidt orthogonalization procedure applied to the basis $b_1, \dots, b_n$ yields the orthogonal vectors $b_i^* = \wp_{n-i+1,L}$, $i = 1, \dots, n$. If n is sufficiently large with respect to c_1 and c then we have $\|b_1\| \|b_n^*\|^{-1} \ge n^{c \log n} c_1^{-c \log c_1}$.*

Theorem 1.9 formulated in Section 1.2 is an immediate consequence of Theorem 4.2.

Proof of Theorem 4.2. In the following lemma we summarize some of the elementary properties of the function $\mathcal{F}_{k,c}$.

Lemma 4.3. *There exist $\alpha_1 > 0, \alpha_2 > 0$ so that if $0 < c < 1$ and k is a positive integer then there is an $\alpha_3 > 0$ so that for all positive integers n, i, $n \ge i$ the following inequalities hold.*

(9). if $i \geq \frac{n}{2}$ then

$$\mathcal{F}_{k,c}(n)(\mathcal{F}_{k,c}(i))^{-1} \leq \alpha_1 n^{2c\log(n/i)};$$

if $i \geq \frac{n}{2} > k$ then

$$\mathcal{F}_{k,c}(n)(\mathcal{F}_{k,c}(i))^{-1} \geq \alpha_2 n^{2c\log(n/i)};$$

(10). for all $\delta > 0$ if $c > 0$ is sufficiently small and $\max\{i,k\} \geq n - 2\log n$ then

$$(1-\delta)\mathcal{F}_{k,c}(n) \leq \mathcal{F}_{k,c}(i);$$

(11). if $0 < \xi < 1$, c is sufficiently small with respect to ξ, n is sufficiently large with respect to c and ξ, $n/2 < i < n - \log n$ then $\mathcal{F}_{k,c}(n)(\mathcal{F}_{k,c}(i))^{-1} \leq (n-i)^{\xi}$.

Proof. (9). α_1 and α_2 are chosen when c and k are already fixed so we may assume that i and n are sufficiently large with respect to c and k. Therefore $\mathcal{F}_{k,c}(n) = \mathrm{e}^{c(\log n)^2}$ and $\mathcal{F}_{k,c}(i) = \mathrm{e}^{c(\log i)^2}$.

$$\begin{aligned} c(\log n)^2 &= c\left(\log i + \log\frac{n}{i}\right)^2 = c(\log i)^2 + 2c\log i \log\frac{n}{i} + c\left(\log\frac{n}{i}\right)^2 \\ &= c(\log i)^2 + 2c\log n \log\frac{n}{i} + 2c(\log i - \log n)\log\frac{n}{i} + c\left(\log\frac{n}{i}\right)^2 . \end{aligned}$$

The assumption $n \geq i \geq n/2$ implies that the absolute values of the last two terms remain below an absolute constant. Therefore the ratio of $\mathcal{F}_{k,c}(n)n^{-2c\log(n/i)}$ and $\mathcal{F}_{k,c}(i)$ remains between two positive constants.

(10). We have

$$\begin{aligned} \mathcal{F}_{k,c}(n-2\log n) &= \exp\left(c(\log(n-2\log n))^2\right) \\ &= \exp\left(c\left(\log\left(n\left(1-\frac{2\log n}{n}\right)\right)\right)^2\right) = \exp\left(c\left(\log n + \log\left(1-\frac{2\log n}{n}\right)\right)^2\right) \\ &= \mathrm{e}^{c(\log n)^2}\exp\left(2c\log n\log\left(1-\frac{2\log n}{n}\right) + \left(\log\left(1-\frac{2\log n}{n}\right)\right)^2\right) . \end{aligned}$$

The assumption that $c > 0$ is sufficiently small implies that we may assume that n is sufficiently large. Since $|\log(1-(2\log n)/n)| < 2(2/n)\log n$ the exponent in the second factor remains below ε in absolute value for any $\varepsilon > 0$ if n is sufficiently large and the first factor is $\mathcal{F}_{k,c}(n)$. Since $\mathcal{F}_{k,c}(i)$ is monotone in the interval $[n-2\log n, n]$ this implies (10).

(11). Inequality (9) implies that there is an absolute constant $c_2 > 0$ so that

$$\mathcal{F}_{k,c}(n)(\mathcal{F}_{k,c}(i))^{-1} \leq c_2 n^{2c\log(n/i)} .$$

Therefore we have to prove that $(n-i)^{\xi}c_2^{-1}n^{-2c\log(n/i)} \geq 1$ provided that $n/2 \leq i < n - \log n$, $c > 0$ is sufficiently small with respect to ξ and n is sufficiently large with respect to c and ξ. Taking the logarithm of both sides of the inequality we get

$$\xi\log(n-i) - \log c_2 - 2c\log n(\log n - \log i) \geq 0 .$$

We consider the left hand side as a function of i. Let $f(x) = \xi \log(n-x) - \log c_2 - 2c\log n(\log n - \log x)$. The derivative of f is

$$f'(x) = -\frac{\xi}{n-x} + \frac{2c\log n}{x}.$$

The function $f'(x)$ is continuous on the interval $[n/2, n-1]$ and it has a single root

$$x_0 = \frac{n}{1+\xi/(2c\log n)}$$

in it (for this latter fact we use that n is sufficiently large with respect to c). Therefore $f'(n/2) > 0$ and $f'(n-1) < 0$ implies that the function f is increasing from $n/2$ till x_0 and then it is decreasing from x_0 till $n-1$. So we can prove our inequality $\xi\log(n-i) - \log c_2 - 2c\log n(\log n - \log i) \geq 0$ for all $i \in [n/2, n-\log n]$ by checking it at the endpoints. If $i = n/2$ then $\log n - \log i \leq \log 2$ and $\log(n-i) \geq (1/2)\log n$, therefore the assumption that c is sufficiently small with respect to ξ implies the inequality. If $i = n - \log n$ then $\log n - \log i < 2n^{-1}\log n$ so the term containing $(\log n - \log i)$ is negligible compared to $\xi\log(n-i) \geq \xi\log\log n$. $\square$

In this proof we will write $\mathcal{F}$ for $\mathcal{F}_{c_1,c}$. First we estimate the probability of the following event A: "$\wp_{n,L}$ is not a shortest nonzero vector in L." Let let $a = (a_1, \ldots, a_n)$ be a sequence of integers. We estimate the probability of the event A_a: "there is an $x \in L$, $x \neq 0$ so that $\|x\| < \|\wp_n\|$ and $\vartheta^{(x,L)} = a$." Let $\texttt{start}(a) = i$. Then A_a implies that $i < n$ (otherwise we would have $x = 0$). We also have $\pi^{(i)}(x) = 0$, $\pi^{(i+1)}(x) \neq 0$. We distinguish four cases depending on i; in each case, we shall estimate $P(A_a)$.

Case I. $\max\{i, c_1\} > n - (\log n)^2$. We show that in this case $P(A_a) = 0$ since there is no $x \in L$ with $\|x\| < \|\wp_n\|$ and $\vartheta^{x,L} = a$. Indeed assume that there is an $x \in L$ with these properties. $\texttt{start}(a) = i$ implies $a_1 = \cdots = a_i = 0$ and so according to the definition of $\vartheta^{(x,L)} = a$ we have $\pi^{(j)}(x) = 0$ for $j = 1, \ldots, i$. $a_{i+1} \neq 0$ and $\pi^{(i+1)}(x) = a_{i+1}\pi^{(i+1)}\wp_{i+1,L} = (0, \ldots, 0, a_{i+1}\mathcal{F}(i+1))$. If $i = n-1$ this implies $\|x\| \geq \wp_{n,L}$ in contradiction to our assumption. Therefore $i < n-1$. (5) implies that

$$\pi^{(i+2)}(x) = \mathcal{U}_{i+1}(a_{i+1}\pi^{(i+1)}\wp_{i,L}) = (0, \ldots, 0, a_{i+1}\mathcal{F}(i+1), \frac{1}{2}\mathcal{F}(i+2)).$$

Consequently

$$\|x\|^2 \geq \|\pi^{(i+2)}x\|^2 \geq (\mathcal{F}(i+1))^2 + \frac{1}{4}(\mathcal{F}(i+2))^2.$$

So the assumption $\max\{i, c_1\} > n - (\log n)^2$ and (10) implies that $\|x\|^2 \geq (\mathcal{F}(n))^2 = \|\wp_{n,L}\|^2$ in contradiction to our indirect assumption.

Case II. $n - (\log n)^2 \geq \max\{i, c_1\}$ and $|S_a| > (1/10)|n - i'|$ where $i' = \max\{i, \lceil n/2 \rceil\}$ and $S_a = \{j \in [i', n] \mid a_j \neq 0\}$. In this case we will use the following inequality:

(12)**.** for all $\gamma > 0$, for all c with $0 < c < 1/20$, for all sufficiently large n and for all x with $n/2 \leq x < n - (\log n)^2$ we have $(n-x)n^{-4c\log(n/x)} > \gamma$.

Taking the logarithm of the left hand side of the inequality we get $\ell = \log(n-x) - 4c\log n(\log n - \log x)$. Since $x \geq n/2$ we have $\log n - \log x = \int_x^n 1/y\, dy \leq \int_x^n 1/(n/2)\, dy = 2n^{-1}(n-x)$. This implies that $\ell \geq \log(n-x) - 8cn^{-1}(n-x)\log n$. If $n - x \leq \sqrt{n}$ then the assumption $x < n - (\log n)^2$ implies that $\ell \geq \log((\log n)^2) - 8cn^{-1}n^{1/2}\log n \geq \log\gamma$ if n is sufficiently large.

If $n-x \geq \sqrt{n}$ then $\ell \geq (1/2)\log n - 8cn^{-1}n\log n \geq (1/2)\log n - 8c\log n$. Since $c < 1/20$ this implies that $\ell \geq \log\gamma$ which completes the proof of (12).

We return now to the discussion of Case II. We show that in this case again $P(A_a) = 0$ since there is no $x \in L$ with $\vartheta^{x,L} = a$ and $\|x\| < \|\wp_n\|$. Indeed assume that there is an $x \in L$ with these properties. By Lemma 3.10 and Lemma 4.3 we have

$$\begin{aligned}\|x\|^2 &\geq \frac{1}{4}\sum\{\|\wp_j\|^2 \mid a_j \neq 0, j \in [i',n]\} \geq \frac{1}{4}|S_x| \cdot \|\wp_{i'}\|^2 \geq |S_x|(\mathcal{F}(i'))^2 \\ &> \frac{1}{4}\alpha_1^2\frac{1}{10}(n-i')n^{-4c\log(n/i')}(\mathcal{F}(n))^2 = \alpha_1^2\frac{1}{40}(n-i')n^{-4c\log(n/i')}\|\wp_n\|^2.\end{aligned}$$

Since $i' \geq n/2$ and c is sufficiently small, inequality (12) is applicable with the substitutions $x := i'$, $\gamma := 40\alpha_1^{-2}$ and we get $\|x\|^2 \geq \|\wp_n\|^2$ in contradiction to our indirect assumption. (Our assumption $n-(\log n)^2 \geq \max\{i,c_1\}$ implies that $n > c_1$. c_1 is sufficiently large with respect to c by the assumptions of Theorem 4.2, and so n is also sufficiently large with respect to c as required in (12).)

Case III. $n-(\log n)^2 \geq \max\{i,c_1\}$, $|S_a| \leq (n-i')/10$ and $i > \lfloor n/2 \rfloor$. (The last inequality implies $i' = i$.) We estimate $P(a)$ in terms of i.

Suppose that L and a are fixed so that there is an $x \in L$ with $\vartheta^{(x,L)} = a$. By the definition of $\vartheta^{(x,L)}$, x is unique with this property. Moreover if we follow the recursive definition of the randomization of $L = \mathtt{rand}_{n,\mathcal{F}}$ then in each step together with $\pi^{(j)}(L)$ we also get $\pi^j(x)$ since by the definition of $\vartheta^{x,L}$ we have $\pi^{(j)}(x) = (\mathcal{U}_{j-1,L}(\pi^{(j-1)})x) + a_j\wp_j$.

Let $Q = [i+2, n-1]\backslash S_a$. If $j \in Q$ then $\vartheta_j^{(x,L)} = a_j = 0$ and so $\pi^{(j)}(x) = \mathcal{U}_{j-1,L}(\pi^{(j-1)}x)$ and therefore by the definition of $\mathcal{U}_{j-1,L}$ we have that $\pi^{(j)}(x) = \pi^{(j-1)}(x) + \rho_j\wp_j$ for a uniquely defined $\rho_j \in (-1/2, 1/2]$. Let $M = \mathcal{F}(n)(\mathcal{F}(i))^{-1}$ and let A'_a be the event that "the number of integers $j \in Q$ with $|\rho_j| < (n-i)^{-1/5}M$ is at least $(n-i)/2$." We show that A_a implies A'_a and therefore $P(A_a) \leq P(A'_a)$, then we will prove an upper bound on $P(A')$. Assume that $\neg A'_a$. Then there are at least $4(n-i)/10$ integers $j \in Q$ so that $|\rho_j| \geq (n-i)^{-1/5}M$. Therefore

$$\begin{aligned}\|x\|^2 &\geq \sum_{j\in Q}\rho_j^2\|\wp_j\|^2 \geq \frac{4}{10}(n-i)(n-i)^{-2/5}M^2(\mathcal{F}(i))^2 \\ &\geq \frac{4}{10}(n-i)^{3/5}\left(\mathcal{F}(n)(\mathcal{F}(i))^{-1}\right)^2(\mathcal{F}(i))^2 \geq \frac{4}{10}(n-i)^{3/5}(\mathcal{F}(n))^2 \geq \wp_n^2\end{aligned}$$

so we reached a contradiction. Therefore $\neg A_a$ holds, which completes the proof of the implication $A'_a \Rightarrow A_a$.

Now we estimate $P(A'_a)$. Suppose $j \in Q$. ρ_j was defined by the conditions $\pi^{(j)}(x) = \pi^{(j-1)}(x) + \rho_j\wp_j$, $\rho_j \in (-1/2, 1/2]$. Therefore the definitions of $\mathtt{rand}$, $\mathtt{ext}$ and Lemma 2.1 imply that if $\pi^{(j-1)}(x)$ is linearly independent of $\wp_{j-1}$ then ρ_j is uniformly distributed over $[-1/2, 1/2]$. However if $\pi^{(j-1)}(x) \neq 0$ and $\wp_{j-1}$ are dependent then by (3) $\mathtt{start}(a) = j-1$. Since $i = \mathtt{start}(a)$, $Q \subseteq [i+2, n]$ this contradicts to $j \in Q$. Therefore the distribution of ρ_j is uniform over $[-1/2, 1/2]$. The definition of $\mathtt{rand}$ implies that the random variables ρ_j, $j \in Q$ are mutually independent. Therefore the probability that "the number of integers $j \in Q$ with $|\rho_j| < (n-i)^{-1/5}M$ is at least $(n-i)/2$" is no more than $2^{n-i}(2(n-i)^{-1/5}M)^{(n-i)/2}$. We get that

(13). $P(A_a) \leq P(A'_a) \leq 2^{n-i}(2(n-i)^{-1/5}M)^{(n-i)/2}$.

We use here the notion that a sequence $a=(a_1,\dots,a_n)$ is *acceptable* as defined before the statement of Lemma 3.18. In this case we use the definition with $h:=\mathcal{F}$.

For a fixed $i>n/2$, let p_i denote the probability that there is an acceptable sequence a with $\mathtt{start}(a)=i$ so that A_a holds. By the union bound, $p_i\le\sum\{P(A_a)\mid a \text{ is acceptable and } \mathtt{start}(a)=i\}$.

$$p_{i,}\le 2^{n-i}(2(n-i)^{-1/5}M)^{(n-i)/2}3^{n-i}\prod_{j=i}^{n}(\mathcal{F}(n)(\mathcal{F}(j))^{-1}).$$

Using that for all $j\in[i,n]$ we have

$$\mathcal{F}(n)(\mathcal{F}(j))^{-1}\le\mathcal{F}(n)(\mathcal{F}(i))^{-1}=M$$

we get that $p_i\le 3^{n-i}2^{3(n-i)/2}M^{3(n-i)/2}(n-i)^{-(n-i)/10}$. By (11) $M\le(n-i)^{\xi}$ for a small constant $\xi>0$ so we get $p_i\le(n-i)^{-(n-i)/20}$. Therefore the probability that there is an i so that A_a holds for a suitably chosen a is at most

$$\sum_{i=n/2}^{n-(\log n)^2}(n-i)^{-(n-i)/20}\le\sum_{k=(\log n)^2}^{\infty}k^{-k/20}.$$

We have that $k^{-k/20}/(k+1)^{-(k+1)/20}\ge(k+1)^{1/20}$ therefore the first term in the sum (with $k=(\log n)^2$) is greater than the sum of the others so we get that the probability is smaller than

$$2((\log n)^2)^{-(\log n)^2/20}\le\exp\left(-\frac{(\log n)\log\log n}{11}\right)$$

if n is sufficiently large; this, however, follows from $n-(\log n)^2>c_1$.

Case IV. $n-(\log n)^2\ge\max\{i,c_1\}$, $|S_a|\le(n-i')/10$ and $i\le\lfloor n/2\rfloor$. (The last inequality implies $i'=\lfloor n/2\rfloor$.)

For each acceptable sequence a with $\mathtt{start}(a)\le n/2$ we will give an upper bound on the probability of the event A_a. Then we will estimate the number of acceptable sequences a with A_a.

Suppose that L and a are fixed so that there is an $x\in L$ with $\vartheta^{(x,L)}=a$. We have $\pi^{(j)}(x)=(\mathcal{U}_{j-1,L}(\pi^{(j-1)})x)+a_j\wp_j$ for all $j=0,\dots,n-1$.

Let $Q=[n/2,n-1]\backslash S_a$. If $j\in Q$ then $\vartheta_j^{(x,L)}=a_j=0$ and so $\pi^{(j)}(x)=\mathcal{U}_{j-1,L}(\pi^{(j-1)}x)$ and therefore by the definition of $\mathcal{U}_{j-1,L}$ we have that $\pi^{(j)}(x)=\pi^{(j-1)}(x)+\rho_j\wp_j$ for a uniquely defined $\rho_j\in(-1/2,1/2]$. Let A'_a be the event that "the number of integers $j\in Q$ with $|\rho_j|<n^{-\tau}$ is at least $n/4$," where $\tau<1/4$ is chosen so that c is sufficiently small with respect to τ. (This definition is somewhat different from the corresponding definition in Case III.) We show that A_a implies A'_a and therefore $P(A_a)\le P(A'_a)$, then we will prove an upper bound on $P(A'_a)$. Assume that $\neg A'_a$. Then there are at least $n/5$ integers $j\in Q$ so that $|\rho_j|\ge n^{-\tau}$. Therefore

$$\|x\|^2\ge\sum_{j\in Q}\rho_j^2\|\wp_j\|^2\ge\frac{1}{5}nn^{-2\tau}(\mathcal{F}(n/2))^2=\frac{1}{5}n^{1-2\tau}(\mathcal{F}(n/2))^2.$$

Thus (11) implies that $\|x\|^2\ge(\mathcal{F}(n))^2=\wp_n^2$, that is, $\neg A_a$ holds, which completes the proof of the implication $A'_a\Rightarrow A_a$.

In Case III we have seen that the random variables ρ_j are independent and uniformly distributed over $[-1/2, 1/2]$. Therefore the probability that "the number of integers $j \in Q$ with $|\rho_j| < n^{-\tau}$ is at least $n/4$" is no more than $2^{n/2}(2n^{-\tau})^{n/4}$. We get that

(14). $P(A_a) \le P(A'_a) \le 2^{n/2}(2n^{-\tau})^{n/4}$.

Let p denote the probability that there is an acceptable sequence a with $\texttt{start}(a) \le n/2$ so that A_a holds. Again by the union bound, $p \le \sum\{P(A_a) \mid a \text{ is acceptable and } \texttt{start}(a) \le n/2\}$. We use (14) as an upper bound on $P(A_a)$. The number of acceptable sequences with the given property is not greater than the total number of acceptable sequences and for this we use the upper bound $3^n \prod_{j=1}^{n} \mathcal{F}(n)(\mathcal{F}(j))^{-1}$ provided by Lemma 3.18.

To get an upper bound on this product we need a lower bound on $\prod_{j=1}^{n}(\mathcal{F}(j))^{-1} \ge c_2 \prod_{j=1}^{n} \mathrm{e}^{-c(\log j)^2}$, where $c_2 > 0$ depends only on c and c_1. Using that

$$\int (\log x)^2 dx = x(\log x)^2 - 2x\log x + 2x + C$$

we give an upper bound on $\sum_{j=1}^{n} c(\log j)^2$. Since the function $(\log x)^2$ is monotone we have

$$\begin{aligned}\sum_{j=1}^{n} c(\log j)^2 &\le c\int_1^n (\log x)^2 dx + (\log n)^2 \\ &= c\left(n(\log n)^2 - 2n\log n + 2n - 2 + (\log n)^2\right).\end{aligned}$$

Substituting this inequality in the upper bound on p we get:

$$\begin{aligned}p &\le 2^{n/2}(2n^{-\tau})^{n/4} 3^n \mathrm{e}^{cn(\log n)^2} \exp\Big(-c\big(n(\log n)^2 - 2n\log n + 2n - 2 + (\log n)^2\big)\Big) \\ &\le 2^{n/2}(2n^{-\tau})^{n/4} 3^n \exp\big(2cn\log n - 2cn - c(\log n)^2 + 2c\big).\end{aligned}$$

Therefore, if $c > 0$ is sufficiently small with respect to τ then $p \le n^{-\tau n/8}$. In other words in Case IV the probability that there is an a so that A_a holds for some suitably chosen a is at most $n^{-\tau n/8}$ assuming $\tau < 1/4$ and c is sufficiently small with respect to τ.

Summarizing the results of the four different cases we get the following:

Lemma 4.4. *If L is the random lattice defined in Theorem 4.2, then the probability that $\wp_{n,L}$ is not a shortest nonzero vector of L is at most $\exp(-(1/12)\log n \log\log n)$. If $n \le c_1$ then $\wp_{n,L}$ is always a shortest nonzero vector.*

Suppose that L is a random value of $\texttt{rand}_{n,\mathcal{F}}$. The definition of the random variable $\texttt{rand}$ implies that $\pi^{(i)}L$ for any $i \le n$ is a random value of $\texttt{rand}_{i,\mathcal{F}}$. Therefore Lemma 4.4 implies that the probability q_n of the event that there is an $i \le n$ such that $\wp_{i,L}$ is not a shortest nonzero vector in $\pi^{(i)}(L)$ is at most

$$\sum_{t=c_1}^{n} \exp\left(-\frac{(\log t)\log\log t}{21}\right).$$

As $n \to \infty$, this series converges (because it is dominated by $\sum_{t=1}^{\infty} t^{-2}$); therefore, if c_1 is sufficiently large then $q_n \le 1/2$. If $i \le c_1$ then $\pi^{(i)}L$ has an orthogonal basis consisting of vectors of the same length so each of them is a shortest nonzero vector. This also implies that $q_n = 0$ for $n \le c_1$. Therefore $q_n \le 1/2$ for all positive integers n. The definition of the basis $\mathbf{b}_{i,L}$, $i = 1, \ldots, n$ implies that $\pi^{(n-i+1)}\mathbf{b}_{i,L} = \wp_{n-i+1,L}$ and so applying the Gram-Schmidt orthogonalization process to it indeed yields the vectors $b_i^* = \wp_{n-i+1,L}$, $i = 1, \ldots, n$. The fact that $\wp_{i,L}$ is a shortest vector in $\pi^{(i)}(L)$ for all $i = 1, \ldots, n$ then implies that $\mathbf{b}_{i,L}$ is a Korkine-Zolotareff basis. The final inequality of the theorem holds since $\|b_1\| = \|\wp_{n,L}\| = \mathcal{F}(n) \ge n^{c \log n}$ and $\|b_n^*\| = \|\wp_{1,L}\| = c_1^{c_1 \log c_1}$. This concludes the proof of Theorem 4.2.

References

[1] * M. AJTAI: Random lattices and a conjectured 0—1 law about their polynomial time computable properties. In *Proc. 43rd FOCS*, pp. 733–742. IEEE Computer Society, 2002. [FOCS:10.1109/SFCS.2002.1181998]. 2

[2] * M. AJTAI: The worst-case behavior of Schnorr's algorithm approximating the shortest nonzero vector in a lattice. In *Proc. 35th STOC*, pp. 396–406. ACM Press, 2003. [STOC:10.1145/780542.780602]. *

[3] * M. AJTAI: Generating hard instances of lattice problems. In J. KRAJIČEK, editor, *Complexity of computations and proofs*, volume 13 of *Quaderni di Matematica*, pp. 1–32. Seconda Universita di Napoli, 2004. Preliminary version: Proc. 28th STOC, 1996, pp. 99–108. 2

[4] * M. AJTAI, R. KUMAR, AND D. SIVAKUMAR: A sieve algorithm for the shortest lattice vector problem. In *Proc. 33rd STOC*, pp. 601–610. ACM Press, 1996. [STOC:10.1145/380752.380857]. 1.1, 1.3, 1.10

[5] * M. L. FURST AND R. KANNAN: Succinct certificates for almost all subset problems. *SIAM Journal on Computing*, 18:550–558, 1989. [SICOMP:10.1137/0218037]. 1.3

[6] * C. F. GAUSS: Recursion der "untersuchungen über die eigenschaften der positiven ternären quadratische formen von ludwig august seeber, dr. der philosophie, ordentl. professor der universität in freiburg, 1831, 248 s. in 4.". *Journal für die reine und angewandte Mathematik*, 20:312–320, 1840. 1.1

[7] * A. M. ODLYZKO J. C. LAGARIAS: Solving low-density subset sum problems. *Journal of the Association for Computing Machinery*, 32(1):229–246, 1985. [JACM:10.1145/2455.2461]. 1.4

[8] * R. KANNAN: Algorithmic geometry of numbers. In *Annual Review of Computer Science*, volume 2, pp. 231–269. Annual Reviews Inc., 1987. 1.1

[9] * R. KANNAN: Minkowski's convex body theorem and integer programming. *Mathematics of Operation Research*, 12(3):415–440, 1987. 1.1, 1.3, 3

[10] * A. KORKINE AND G. ZOLOTAREFF: Sur les formes quadratiques. *Mathematische Annalen*, 6:366–389, 1873. [Springer:p56345710m4p6214]. 2

[11] * J. L. LAGRANGE: Recherches d'arithmétique. In M. J.-A. SERRET, editor, *Oeuvres de Lagrange*, volume 3, pp. 698–701. Gauthier-Villars, 1869. (article cca 1773). 1.1

[12] * A. K. LENSTRA, H. W. LENSTRA, AND L. LOVÁSZ: Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261:515–534, 1982. [Springer:lh1m24436431g068]. 1.1

[13] * A. M. ODLYZKO AND H. TE RIELE: Disproof of the Mertens conjecture. *Journal für die reine und angewandte Mathematik*, 357:138–160, 1985. 1.4

[14] * C.-P. SCHNORR: A hierarchy of polynomial time lattice basis reduction algorithms. *Theoretical Computer Science*, 53:201–224, 1987. [TCS:10.1016/0304-3975(87)90064-8]. 1.1, 1, 3

[15] * C.-P. SCHNORR: Lattice reduction by random sampling and birthday methods. In *Proc. 20th Ann. Symp. on Theoretical Aspects of Computer Science (STACS'03)*, volume 2607 of *Lecture Notes in Computer Science*, pp. 145–156. Springer, 2003. [STACS:qjpadpmwabty52g4]. 1.1

AUTHOR

Miklós Ajtai
IBM Almaden Research Center
ajtai@almaden.ibm.com

ABOUT THE AUTHOR

MIKLÓS AJTAI received his Ph. D. from the Hungarian Academy of Sciences in 1975. His advisor was András Hajnal. He worked in the following areas: axiomatic set theory (independence proofs), lattice theory (posets with meet and join), combinatorics, the theory of random graphs, complexity theory, sorting networks, the theory of lattices (n-dimensional grids) and their applications to complexity theory and cryptography. He is a member of the Hungarian Academy of Sciences and was an invited speaker at ICM in 1998. He received the Knuth prize in 2003, and the IBM Corporate Award in 2000.